

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2004, 19:55:00 ; Search time 408 Seconds

(without alignments)
8496.396 Million cell updates/sec

Title: US-10-089-557-2_COPY_1_816

Perfect score: 816
Sequence: 1 actataggcagcgcgtgctc.....tatttcagtcattgtaaaag 816

Scoring table: IDENTITY NUC
Gapop 10'-0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	816	100.0	3041	6	AAS19999 Cotton CF
2	124.6	15.3	8056	7	ABZ10246 Haematopo
3	123	15.1	8056	7	ABZ10100 Haematopo
4	103.2	12.6	8056	7	ABZ10246 Haematopo
5	97.8	12.0	8056	7	ABZ10100 Haematopo
6	97	11.9	8310	2	AAZ29911 CDNA enco
7	94.2	11.5	494	5	ABV10021 Human pro
8	90.2	11.1	7351	6	ABL32029 Human imm
9	90.2	11.1	8305	6	ABL33569 Human imm
10	89.8	11.0	18357	6	ABG67084 Human ang
11	89.4	11.0	6106	4	AAS46429 Tumour su
12	89.4	11.0	6106	6	ABK40031 Human che
13	89.4	11.0	6106	6	ABL33472 Human imm
14	89	10.9	6478	9	ADB54284 Pretreat
15	88.2	10.8	16602	6	ABL32727 Human imm
16	88.2	10.8	16602	6	ABN80069 Human che
17	88	10.8	5852	2	AAQ11710 Dictyoste
18	87.8	10.7	419	7	ABX46069 Bovine ES
19	87.4	10.7	17848	4	AAS45323 Chemicall
20	87.4	10.7	17848	4	ABK39976 Human che
21	87.4	10.7	17848	6	ABK28164 DNA trans
22	87.2	10.7	15548	6	ABL34155 Human imm
23	86.6	10.6	1501	7	ABZ10188 Haematopo

24	86.6	10.6	1501	9	ADB84162
25	86	10.5	2993	9	ADB37661
26	86	10.5	38342	4	AAS46746
27	86	10.5	38342	6	ABK31507
28	85.8	10.5	5487	6	ABL33598
29	85.8	10.5	7508	6	ABK31207
30	85.4	10.5	3051	7	ACF62818
31	85.4	10.5	5952	9	ADB54320
32	85.2	10.4	6465	6	ABL32985
33	85.2	10.4	9789	2	AAT41852
34	84.8	10.4	1671	2	AAQ24134
35	84.8	10.4	5845	6	ABL33663
36	84.6	10.4	34548	6	ABL70604
37	84.4	10.3	6063	4	AAS46337
38	84.4	10.3	6216	4	ABK39932
39	84.4	10.3	6216	6	ABL70139
40	84.4	10.3	13511	6	ABL32281
41	84.4	10.3	34548	6	ABL70603
42	84.2	10.3	6478	4	AAS45417
43	84.2	10.3	6478	6	ABK28270
44	84.2	10.3	6478	6	ABN80201
45	84.2	10.3	6478	9	ADB54156

ALIGNMENTS

RESULT 1
ID AAS19999 standard; DNA; 3041 BP.
XX AAS19999;
AC AAS19999;
DT 21-MAY-2002 (first entry)
XX
XX
DE Cotton CFACT1 gene sequence.
XX
KM Cotton: fibre-specific actin gene; CFACT1; improved cotton fibre;
KM high fibre quality; cotton yield; transgenic; plant; anthocyanin;
KM coloured cotton; silk protein; silk worm; spider; polyhydroxybutyrate;
KM improved thermal property; insulating characteristic; gene; ds.
XX
OS Gossypium hirsutum.
XX
XX
XX WO200210413-A1.
XX
XX
XX 07-FEB-2002.
XX
XX 01-AUG-2000; 2000WO-SG000112.
XX
XX 01-AUG-2000; 2000WO-SG000112.
XX
XX 01-AUG-2000; 2000WO-SG000112.
XX
XX (MOLE-) INST MOLECULAR AGRICULTURE.
XX
XX Li X, Cai L, Cheng N, Liu J;
XX WPI; 2002-217123/27.
XX
XX
XX New fiber-specific actin promoter from cotton for controlling gene
XX expression in cotton fibers and creating transgenic plants, in particular
XX cotton plants, having altered fiber characteristics.
XX
XX Claim 2; Fig 2; 36pp; English.
XX
XX The present invention relates to the isolation of the cotton fibre-
XX specific actin gene, CFACT1, which includes the promoter sequence at the
XX cotton CFACT1 gene promoter controls specific gene expression at the
XX transcriptional level in cotton fibres. It is useful for improving cotton
XX fibres to create new cotton varieties with high fibre quality and yield,
XX by gene manipulation methodologies. The promoter is also useful for
XX creating transgenic plants, particularly cotton, having altered fibre
XX characteristics, and permits selective expression of a transgene in the
XX cotton fibre, permitting greater latitude in the types of transgenes

CC employed. Examples of expression of desirable genes in cotton fibres, but
 CC not in other parts of the cotton plants, include anthocyanin genes for
 CC coloured cotton, silk protein genes from silk worm or spiders for
 CC increased strength cotton fibres, and biosyntheses of polyhydroxybutyrate
 CC in cotton fibres for improved thermal properties and insulating
 CC characteristics. The present sequence represents the cotton CTRC11 gene
 XX

Sequence 3041 BP; 892 A; 517 C; 629 G; 1003 T; 0 U; 0 Other;

Query Match 100.0%; Score 816; DB 6; Length 3041;
 Best Local Similarity 100.0%; Pred. No. 5e-90;
 Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTATAGGGGACGCGTGTGAGCGCCCGGCTGTCTCTTAAGAAACAATTGTGTCAAG 60
 DB 1 ACTATAGGGGACGCGTGTGAGCGCCCGGCTGTCTCTTAAGAAACAATTGTGTCAAG 60
 QY 61 TCGTTCTTCCGAGCAATCCGAATAGGAGCTTAGAGTAACTCTAACACAGGACTGCT 120
 DB 61 TCGTTCTTCCGAGCAATCCGAATAGGAGCTTAGAGTAACTCTAACACAGGACTGCT 120
 QY 121 CCAGCATTAACCTGTGGTGAATAATGTTATGGTAGTGTGATGCGAAGTATTTGATGG 180
 DB 121 CCAGCATTAACCTGTGGTGAATAATGTTATGGTAGTGTGATGCGAAGTATTTGATGG 180
 QY 181 AAGGTGTGAAGATTAATGTTATGGGATTAATTTCTAGATTAATTTGGTTTGA 240
 DB 181 AAGGTGTGAAGATTAATGTTATGGGATTAATTTCTAGATTAATTTGGTTTGA 240
 QY 241 AGTAT 300
 DB 241 AGTAT 300
 QY 301 TAATTTATTTAATCTTGTGTTTCAATTTATGACGGTTAATATTTTGAATTAAT 360
 DB 301 TAATTTATTTAATCTTGTGTTTCAATTTATGACGGTTAATATTTTGAATTAAT 360
 QY 361 GAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
 DB 361 GAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
 QY 421 ATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 DB 421 ATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 QY 481 AAT 540
 DB 481 AAT 540
 QY 541 TTGCGAGTTCATCCAAACACGAGAAAGTAAATCAATTTTCAAGAAAGTAAATCAAT 600
 DB 541 TTGCGAGTTCATCCAAACACGAGAAAGTAAATCAATTTTCAAGAAAGTAAATCAAT 600
 QY 601 AGAAATTTATTTTGGAAATTAATTTACTGCGCAAAACAATGAGTCAATGTTCTGTT 660
 DB 601 AGAAATTTATTTTGGAAATTAATTTACTGCGCAAAACAATGAGTCAATGTTCTGTT 660
 QY 661 TTTATTTTATTTTCTATTTAGAGAACTAGAAATTAATTTGTCAAATGCTTTAATC 720
 DB 661 TTTATTTTATTTTCTATTTAGAGAACTAGAAATTAATTTGTCAAATGCTTTAATC 720
 QY 721 TAGCTGTTTAAATAGTTGAAGGACAGAAACCCCGGTGTCAGTGTGCTGATC 780
 DB 721 TAGCTGTTTAAATAGTTGAAGGACAGAAACCCCGGTGTCAGTGTGCTGATC 780
 QY 781 TCACCTACCTAGATCTTATTTTCAATTTGTAATTAAG 816
 DB 781 TCACCTACCTAGATCTTATTTTCAATTTGTAATTAAG 816

RESULT 2
 ABZ10246
 ID ABZ10246 standard; DNA; 8056 BP.

XX ABZ10246;
 AC 16-JUN-2003 (first entry)
 XX
 DT
 XX
 DE Haematopoietic cell proliferation disorder related DNA sequence #386.
 XX
 KM Human; haematopoietic cell proliferation disorder; cytostatic;
 KM gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 XX cytosine methylation state; gene; ds.
 OS Homo sapiens.
 PN WO200277272-A2.
 PD 03-OCT-2002.
 PF 26-MAR-2002; 2002WO-EP003401.
 PR 26-MAR-2001; 2001US-0278333P.
 PA (EPig-) EPIGENOMICS AG.
 PI Berlin K, Braun A, Distler J, Gueig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Walter S, Model F, Mueller V, Otto T, Pelet C;
 PI Schwope I, Ziebarth H;
 XX
 DR MPI; 2003-018942/01.

XX
 PT Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.
 PS Claim 28; SEQ ID NO 386; 117bp; English.

XX The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclases, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 XX

Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;
 Query Match 15.3%; Score 124.6; DB 7; Length 8056;
 Best Local Similarity 53.3%; Pred. No. 3.9e-07;
 Matches 285; Conservative 0; Mismatches 249; Indels 1; Gaps 1;

QY 189 AAGATTAATGTTATTTGGATTAATTTCTAGATTAATTTGCTTGGAAAGTAAATA 248
 DB 1646 AAAAAATTTATTTAATTAATTAATTAATTTATTTTATTTAATTAATTAATTAAT 1705
 QY 249 TATATATTAATCAATCTGTTTATTTTATTTTATTAACAAATTAATTAATTAAT 308
 DB 1706 TATATATTAATTAATTAATTTATTTATTTATTTAATTAATTAATTAATTAATTT 1765

QY 309 TTACTTTGGTTTTCATTTATGACGGT-TAATATTTTACGTTTAAATGAGCATT 367
DB 1766 TTAATATTAATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTA 1825
QY 368 ATATATATTAATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTA 427
DB 1826 ATTTTGTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1885
QY 428 TATTAATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 487
DB 1886 TTTTATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1945
QY 488 AAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 547
DB 1946 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2005
QY 548 TTCAATCAACACACAGAAAGTAATATCAATTTTCAGAAAGTAATCAATTTTCAGAAAT 607
DB 2006 AAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2065
QY 608 ATTTTTCGGAATTAATTTTACGCAACAAATGAGCTAAGTTTCTGTTTAAATT 667
DB 2066 TTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2125
QY 668 TTTATTTTCTATTTAGAGAACTAGAAATGATTTGCAAAATGCTTTAACTTA 722
DB 2126 TAAATTTTATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 2180

RESULT 3
ABZ10100
ID ABZ10100 standard; DNA; 8056 BP.
XX ABZ10100;
AC
XX 16-JAN-2003 (first entry)
DT
XX Haematopoietic cell proliferation disorder related DNA sequence #240.
DE
XX Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
XX WO200277272-A2.
XX
XX 03-OCT-2002.
XX
XX 26-MAR-2002; 2002WO-EP003401.
XX
XX 26-MAR-2001; 2001US-0278333P.
XX
XX (EPIC-) EPICENOMICS AG.
XX
XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Pienbrock C, Adorjan P, Grabs G, Lesche R, Leu E; Pelet C;
PI Levin A, Lipscher B, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwobe I, Ziebarth H;
XX
XX WPI; 2003-018942/01.
XX
XX Detecting and differentiating between haematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
XX Claim 28; SEQ ID NO 240; 117bp; English.

CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferation disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX

SO Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;

Query Match 15.1%; Score 123; DB 7; Length 8056;
Best Local Similarity 53.1%; Pred. No. 6.1e-07;
Matches 284; Conservative 0; Mismatches 250; Indels 1; Gaps 1;

QY 189 AAGATTAATGTAATGGGATTACTAATTTCTAGTATTAATGCGTTGGAGTTATA 248
DB 1646 AAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1705
QY 249 TATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 308
DB 1706 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1765
QY 309 TTAATTTGTTGTTTCAATTTATGACGGT-TAATATTTTATGTTTAAATTAATGACATT 367
DB 1766 TTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1825
QY 368 ATATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 427
DB 1826 ATTTGTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1885
QY 428 TATTAATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 487
DB 1886 TTTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1945
QY 488 AAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 547
DB 1946 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2005
QY 548 TTCAATCAACACACAGAAAGTAATATCAATTTTCAGAAAGTAATCAATTTTCAGAAAT 607
DB 2006 AAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2065
QY 608 ATTTTTCGGAATTAATTTTACGCAACAAATGAGCTAAGTTTCTGTTTAAATT 667
DB 2066 TTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2125
QY 668 TTTATTTTCTATTTAGAGAACTAGAAATGATTTGCAAAATGCTTTAACTTA 722
DB 2126 TAAATTTTATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 2180

RESULT 4
ABZ10246/c
ID ABZ10246 standard; DNA; 8056 BP.
XX ABZ10246;
AC
XX 16-JAN-2003 (first entry)
DT
XX Haematopoietic cell proliferation disorder related DNA sequence #386.
DE
XX Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;

CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients

XX Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;

Query Match 12.0%; Score 97.8; DB 7; Length 8056;

Best Local Similarity 49.6%; Pred. No. 0.00065;

Matches 284; Conservative 0; Mismatches 282; Indels 7; Gaps 1;

QY 170 TATTTTCATGAGAGGTGTTAAGAAATTAATGTTATGAGATTAATTTCTAGATTAAAT 229

Db 2644 TTTTATTTATTTTATTTTATTTTATTTATTTATTTATTTTATTTTATTTTATTTT 2585

QY 230 TGTGCTTGGAGTTAATATATATATATCAATCCTGTTTATTTTATTTTATTTTATTT 289

Db 2584 TATTTTAAAAAATTTATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 2525

QY 290 ACAATTCAGAAATATTTATTTTAACTTGCTGTTTCAATTTATGACGGTAAATATTTTA 349

Db 2524 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2465

QY 350 GTTAAATATATGACATATATATATATATTAATTAATTAATTAATTAATTAATTAATTA 409

Db 2464 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2405

QY 410 TAAATTTAAATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 469

Db 2404 TCGAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2345

QY 470 ATAAATCTTAAATATTAATATTTTAAATAGCTTTCAGTAAATCTGCAACAATA 529

Db 2344 AAATTAATAAATTAATAAATTAAGAAAAATTAATAAATTAATAAATTAATAAATTAATA 2285

QY 530 GAAATATATTTTGGAGGTGATCCAAACACGAGAAAGTAATCATTTTCAGAAAGTA 589

Db 2284 AATTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 2225

QY 590 AATCATTTTTCAGAAATTAATTTTTCGAAATTAATTTTCTGCAAAACAAATGGAGTCTAA 649

Db 2224 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2169

QY 650 GGTGTTCTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 709

Db 2168 ---TT 2112

QY 710 TGTCTTAATCTAGCTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 742

Db 2111 AATAATTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2079

RESULT 6

AAZ29911/C

ID AAZ29911 standard; DNA; 8310 BP.

XX AAZ29911;

XX 26-JAN-2000 (first entry)

XX CDNA encoding a SC4 protein of soybean.

XX Promoter; seed coat specificity; SC4 protein; herbicide resistance gene;

XX viral coat protein; biological control; Bt toxin; seed taste; ss.

OS Glycine max.

XX Key Location/Qualifiers

FT promoter 1..5514

FT /note="Claimed"

PN W09953067-A2.

PD 21-OCT-1999.

PF 13-APR-1999; 99WO-CA000293.

PR 13-APR-1998; 98US-00059090.

PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

PI Miki B, Gijzen M, Miller S, Bowman L, Batchelor A, Hu M;

PI Boulhiser K;

PS WPI; 1999-611304/52.

PT Novel promoter sequences and genes useful for inducing expression of

PT genes in plant seed coats.

PS Claim 36; Page 147-155; 155pp; English.

CC The present sequence encodes a SC4 protein. The promoter of this gene is

CC differentially expressed in seed coat tissues, specifically within the

CC outer integument, the inner integument, the thick walled parenchyma, the

CC thin walled parenchyma, the endosperm, the houghless cells, the

CC palisade, the stellate parenchyma, or the membranous endocarp associated

CC with the seed coat. The seed-coat promoters may be used to express

CC proteins of interest in seed coat tissues. Genes of interest include but

CC are not restricted to herbicide resistance genes, genes encoding viral

CC coat proteins, or genes encoding proteins conferring biological control

CC of pests or pathogens, e.g. a Bt toxin. Other genes that may be expressed

CC include proteins that alter the taste of the seed and/or affect the

CC nutritive value of the seed

XX Sequence 8310 BP; 2807 A; 1283 C; 1265 G; 2955 T; 0 U; 0 Other;

Query Match 11.9%; Score 97; DB 2; Length 8310;

Best Local Similarity 48.9%; Pred. No. 0.0008;

Matches 291; Conservative 0; Mismatches 300; Indels 4; Gaps 1;

QY 160 TATGCGAGTATTTTCATGGAAGGTGTTAAGAAATTAATGTTATGGAATTAATTTTC 219

Db 4944 TTTTAAATTTATTAATAAATTTATGAATTAATAAATAAATTTATTAATATCCTTTTATA 4885

QY 220 TAGTATTAATTTGCTGTTGGAAGTTAATATTAATTTATTTCAATCCTGTTTATTTT 279

Db 4884 AATCCTTTAATTTTCTGTATGTCATTAATTAATTAATTAATTAATTAATTAATTAATA 4825

QY 280 TTTTATTAACCAATTAACAATTAATTTATTTAACTTGCTGTTTCAATTTATGACGCT 339

Db 4824 TATATATATATATATATATGAAATGTTATATATTTTATTTTATTTTATTTTATTTAT 4765

QY 340 TAAATTTTATGTTTAAATTAATTTGAGCATTAATTAATTAATTAATTAATTAATTAATTA 399

Db 4764 TAAATTTTATGTTTAAATTAATTTGAGCATTAATTAATTAATTAATTAATTAATTAATTA 4705

QY 400 TATGTAATAATTAATTTAAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTT 459

Db 4704 TGT---AAATTAACCTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4649

QY 460 TAAAGATTAATTAATTTCTTAATTAATTTTATTTTAAATTAATTTTAAATTAATTTTAAATTT 519

Db 4648 TTTAGAAAAAACAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4589

QY 520 TCAACAATTAAGAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 579

Db 4588 TTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4529

QY 580 CAGAAAGTAATCATTTTCAGAAATTTTCGAAATTTTCTAGCGCAACAA 639
 Db 4528 TATATTAATTTATTTATGATATGATAAAATTCATCATTTATTTTATATTAATTA 4469
 QY 640 TGGAGCTAGAGTCTGCTGTTTTATTTTATTTTCTATTTAGAGAACTGAAATTC 639
 Db 4468 TGAAGAAAGTTATTTTATTTATTAACAAGTAAATTAATTTTATGATTTAGTTAATG 4409
 QY 700 ATTTGCTAAATGCTTTAATCTAGCTGTTTATGATTGTAAGGACAGAAAC 754
 Db 4408 TTTATTTAAATCTACTTAATTTTCATTAATTAAGTAAATTAATTTTAAATTAATC 4354

RESULT 7

ABV10021/c
 ID ABV10021 standard; cDNA; 494 BP.

AC ABV10021;

DT 13-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 10012.

KW Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US005171.

PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI: 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of
 prostate cells and correlating with presence of prostate cancer, useful
 for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 1602; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 specification or its complement. (I) is useful for: (a) assessing whether
 a patient is afflicted with prostate cancer; (b) monitoring the
 progression of prostate cancer in a patient; (c) assessing the efficacy
 of a test compound to inhibit prostate cancer in a patient; (d) assessing
 the efficacy of a therapy for inhibiting prostate cancer in a patient;
 (e) selecting a composition for inhibiting prostate cancer in a patient;
 (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 determining whether prostate cancer has metastasized in a patient; (h)
 assessing the aggressiveness or indolence of prostate cancer in a patient
 ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 494 BP; 299 A; 19 C; 21 G; 155 T; 0 U; 0 Other;

Query Match 11.5%; Score 94.2; DB 5; Length 494;

Best Local Similarity 57.2%; Pred. No. 0.0029;
 Matches 194; Conservative 0; Mismatches 138; Indels 7; Gaps 1;

QY 134 TTTGGTAAATGCTATGATGCTAGTGCAGATTTTCATGGAAGGTGTAAGAA 193
 Db 405 TTTTATTAATTTTATTTATTAATTTTATTTTATTTATTTATTTATTTATTTATTA 346
 QY 194 TTAATGTTATTTGGATTCTAATTTCTAGTATTATTTGGTTGGAAGTTAATAA 253
 Db 345 TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 286
 QY 254 TTAATCAATCTGTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 313
 Db 285 TATT-----TTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 233
 QY 314 TTTGTTGTTTCAATTTATGACGTTAATTTATTTAGTTAATTTATTTATTTATTTATA 373
 Db 232 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTA 173
 QY 374 TATTAAATTAATTAATCAATTTGTAATTAATGTAATAATTTAAATTAATTTATTA 433
 Db 172 TATTAATATTAATAAATAATTAATTTATTAATTAATAAATAAATAAATAAATAAATAA 113
 QY 434 TATATATTAATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 472
 Db 112 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 74

RESULT 8

ABL32029
 ID ABL32029 standard; DNA; 7351 BP.

AC ABL32029;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 2.

XX Human; immune system disease; cytosine methylation; antiaesthetic;
 KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 ds.

OS Homo sapiens.

PN WO200200928-A2.

PD 03-JAN-2002.

PF 02-JUL-2001; 2001WO-EP007537.

PR 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

PA (EPIC-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI: 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful for
 diagnosis and treatment of diseases associated with abnormal cytosine
 methylation.

PS Claim 1; SEQ ID NO 2; 32pp + Sequence listing; German.

XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC Rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 XX
 SQ Sequence 7351 BP; 2433 A; 42 C; 1458 G; 3418 T; 0 U; 0 Other;

Query Match 11.1%; Score 90.2; DB 6; Length 7351;
 Best Local Similarity 49.4%; Pred. No. 0.0054;
 Matches 293; Conservative 0; Mismatches 293; Indels 9; Gaps 2;

QY 140 GAAATGTTAATGCTGATGTCGAGTATTTTCATGAGGAGTGTAGAAATTAAG 199
 DB 897 GATTAATTTTATATTTTATTTTATGATTAATTAATTTGATGATTAATTAATTA 956
 QY 200 TTAATGAGATTTACTAATTTCTAGTATTAATTTGTTGGAGCTAATATATTAATTC 259
 DB 957 TTAATTAAGATATTAATAAATTTTAAATGTTGATTAATTAATTTTATTTT 1016
 QY 260 AATCCTGTTTATTTTATTTTATTAACACATTAACAAATATTTTAACTTGGT 319
 DB 1017 TTTTATTTTATTTGAAATTTGAAATTTGATTAATTAATTAATTAATTAATTA 1076
 QY 320 TGTTCATTAATTAATGAGCTTAATTTTATTAATTAATTAATTAATTAATTA 379
 DB 1077 TTTTAATTTTAAATAATTTGGAATTTATTTTATTTTATTTTATTTTATTTAT 1136
 QY 380 ATAAATTAATCATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 439
 DB 1137 ATATAGCTAT-----GATATTTATGTTTATTAATTAATTAATTAATTAATTA 1189
 QY 440 TAATAAATCTCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 499
 DB 1190 TAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1249
 QY 500 ATACCTTTGAGTAAATCTGTCAGAACATTAATTAATTTTTCAGGTTTCATCAACA 559
 DB 1250 ATCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1309
 QY 560 CCAGAAAGTAAATCATTTTCAGAAAGTAAATCATTTTCAGAA--ATTATTTTCGGA 617
 DB 1310 AAAATATTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1369
 QY 618 AATTAATTTTACTGCAAAATTAATTAATTAATTAATTAATTAATTAATTTTTC 677
 DB 1370 AGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1429
 QY 678 TATTAGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 734
 DB 1430 TTTTATGGAATTTATGATTTATGATTTGATTAATTAATTAATTAATTTTGT 1486

RESULT 9
 ABL33569
 ID ABL33569 standard; DNA; 8305 BP.

XX ABL33569;
 AC
 DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1542.

XX Human; immune system disease; cytosine methylation; antiaesthetic;
 KW antiaesthetic; antiaesthetic; antiaesthetic; antiaesthetic;
 KW antiaesthetic; antiaesthetic; antiaesthetic; antiaesthetic;
 KW antiaesthetic; antiaesthetic; antiaesthetic; antiaesthetic;
 KW antiaesthetic; antiaesthetic; antiaesthetic; antiaesthetic;
 KW acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 ds.

OS Homo sapiens.
 XX
 XX WO200200928-A2.
 XX

PD 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP007537.

XX 30-JUN-2000; 2000DE-01032529.

XX 01-SEP-2000; 2000DE-01043826.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.

XX Claim 1; SEQ ID NO 1542; 32pp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention

SQ Sequence 8305 BP; 2163 A; 185 C; 2060 G; 3897 T; 0 U; 0 Other;

Query Match 11.1%; Score 90.2; DB 6; Length 8305;
 Best Local Similarity 50.1%; Pred. No. 0.0052;
 Matches 284; Conservative 0; Mismatches 273; Indels 10; Gaps 2;

QY 173 TTTGATGAAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 232
 DB 1710 TTTTATGTTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1769
 QY 233 GGTTCGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1829
 DB 1770 ATATAGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1889
 QY 293 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 352
 DB 1830 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1889
 QY 353 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 412
 DB 1890 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1949
 QY 413 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 472
 DB 1950 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2003
 QY 473 AATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 532
 DB 2004 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2063
 QY 533 AATATTTTTCAGGTTGATCAACACACAGAAAGTAAATCATTTTCAGAAAGTAAAT 592
 DB 2064 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2119
 QY 593 CATTTTTCAGAAATTAATTTTCGAAATTAATTTTCGCAAAATTAATTAATTAATTA 652
 DB 2120 TATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2179
 QY 653 TTTGCTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 712
 DB 2180 TATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2239
 QY 713 CTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 739
 DB 2240 ATATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2266

Db	15466	TTTTTTTATTTTGGTATTATATAGGGTAGTTATGTTTGGTTTAAATTAAGTTGGAG	15525
Qy	364	CATTCTTATATATTAATTAATAATCATTTGTATATATGTAATAAATTAATTAATAATA	423
Db	15526	GTGTTTTAAGAAATTTTAAAGAGATTTGTAATAATTAATTAATTAATTAATAAGTAA	15585
Qy	424	AATTAATTAATATATATTAATTAATCTCATCAACAAATTAATAAGATAAATCTTAAT	483
Db	15586	GATTATTTTAAATGTTTTTTTTTATTTATATATAGTGGAATTTCCATTATTTTTTTATAT	15645
Qy	484	ATATATAATTTTTTAAAAATGCTTTTCAGTAATCTGTCAACAATAGAAAATATTTTTTG	543
Db	15646	AATATAGT-----AATAGATGAAAATTAATAATATTAATATATATAGTAATTTTAT	15698
Qy	544	CAGTCTATCCAAACACAGAAAAGTAATCATTTTTCAGAAAAGTAAATCATTTTTCAGA	603
Db	15698	TATGTATATTTTAAAGAGAGATTTAGACTTATTTTTTGATAGTAATTAATTTTTTTTAT	15758
Qy	604	AATTAATTTTTCGAAATTAATTTTACTGCGAAACAAAT-GGAGCTAAGTGTTCTGTTT	662
Db	15759	TTTTATTTGTTTTTTTAAATTTTATTTTATTTTATTAATTAAGTTTAAAGTTTGTGT	15818
Qy	663	TATTTTATTTTTCATTTTAAAGAACTAGAAATGATTTGTCAAAATGCTTTAATCTA	722
Db	15819	TATTTTGGTTTTTATTTATTTTGTGTTTTTGTGTTTTTGTGTTTTTATTAATTTTTT	15878
Qy	723	GCTTGTTAGATTAGTT 739	
Db	15879	TTTTTATTTAGTTTTTT 15895	
RESULT 11			
AA	46429		
ID	AA	46429 standard; DNA; 6106 BP.	
AC	AA	46429;	
XX	XX		
DT	18-DEC-2001	(first entry)	
XX	XX		
DE	Tumour suppressor gene derived chemically modified sequence #151.		
XX	XX		
KW	Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;		
KW	tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;		
XX	cytosine methylation; ds.		
OS	Homo sapiens.		
XX	XX		
PN	WO200168912-A2.		
XX	XX		
PD	20-SEP-2001.		
XX	XX		
PF	15-MAR-2001; 2001WO-EP002955.		
XX	XX		
PR	15-MAR-2000; 2000DE-01013847.		
PR	06-APR-2000; 2000DE-01019058.		
PR	07-APR-2000; 2000DE-01019173.		
PR	30-JUN-2000; 2000DE-01032523.		
PR	01-SEP-2000; 2000DE-01043826.		
XX	XX		
PA	(EPIC-) EPIGENOMICS AG.		
XX	XX		
PI	Olek A, Piepenbrock C, Berlin K;		
XX	XX		
DR	WPI; 2001-602752/68.		
XX	XX		
PT	Fragments of chemically modified genes associated with tumor suppressor		
PT	genes and oncogenes, useful in designing primers and probes for analyzing		
XX	diseases associated with cytosine methylation state e.g. cancer.		
PS	Claim 1; SEQ ID NO 151; 27pp; English.		
CC	The invention relates to a nucleic acid comprising a sequence of 18		
CC	bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with		
CC			

bisulphite, of genes associated with tumour suppression and oncogenes having a sequence taken from 536 (actually 533 since numbers 408, 458 and 500 are missing from the sequence listing) sequences (SS) and sequences complementary to (SS). The nucleic acid may be a peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may form part of a set of probes for detecting the cytosine methylation state and/or single nucleotide polymorphisms and also to be used in an array for analysing diseases associated with CpG dinucleotides e.g. cancers and tumours. The probes can also be used in a method for ascertaining genetic and/or epigenetic parameters for the diagnosis and/or therapy of existing diseases or the predisposition to specific diseases, by analysing cytosine methylations. The parameters may be compared to another set of genetic and/or epigenetic parameters, the differences serving as basis for diagnosis and/or prognosis events which are disadvantageous to patients. The present sequence is one of the 533 genomic sequences derived from tumour suppressor genes and oncogenes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 6106 BP; 1938 A; 30 C; 849 G; 3289 T; 0 U; 0 Other;

Query Match 11.0%; Score 89.4; DB 4; Length 6106;
Best Local Similarity 51.7%; Pred. No. 0.0069;
Matches 278; Conservative 0; Mismatches 253; Indels 7; Gaps 3;

```

OY 198 TGTATGCGATTACTAATTTCTAGTATTAATGTTGGTTGGAAGTTAATATTAAT 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2612 TTTTATTTTAAATGTTTTGTTGTTTAAAGTTAAAGAAATTTAAATAGTTA 2671

OY 258 TCAATCCTGTTGTTTATTTTATTTTAAACAAATTAACAATTTATTTACTTG 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2672 TTAATGTTTAAATTTGTTTAAATTAATTAATTAATTAATTAATTAATGTTA 2731

OY 318 GTTGTTTCAATTTATGACGTTAATTTTATGTTTAAATAGACATTTATATAT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2732 TATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2791

OY 378 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2792 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2850

OY 438 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2851 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2910

OY 493 TTTTAAATTAATGCTTTTCAATTAATGTCGCAACAATTAATTAATTTTGGAGTTCAT 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2911 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2970

OY 553 CCAAGACCCAGAAAGTAATCAATTTTCAGAAAAGTAATCAATTTTCAGAAATTAATT 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2971 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3029

OY 613 TCGAAATTAATTTTAACTGCAACAATGAGCTTAAGTGTTCGTTTATTTTAT 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3030 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 3089

OY 673 TTTTCTATTTAGAAACTAGAAATGATTTGTCAAAGCTTTTATCTAGCTGTT 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3090 TATAGTATTTATTAATTAATTTGTTGTTTATGAGCTTATGTAATTTGTTT 3147

```

RESULT 12

ABK40031 standard; DNA; 6106 BP.

ABK40031;

21-MAY-2002 (first entry)

Human chemically pretreated gene sequence #57 strand 1.

XX

```

KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KM cytotoxic; ALDH6; CYP11A; CYP11B; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW UGT; MRP; pharmacogenetics; SNP; single nucleotide polymorphism.
OS Homo sapiens.
PN MO200202806-A2.
PD 10-JAN-2002.
XX 29-JUN-2001; 2001WO-EP007470.
XX 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX (EPIC-) EPIGENOMICS AG.
PA Olek A, Piepenbrock C, Berlin K;
PI WPI; 2002-154757/20.
DR New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT useful for detecting cytosine methylation state of genes associated with
PS pharmacogenetics and for therapy of diseases e.g. cancer.
PS Claim 1; SEQ ID NO 113; 24pp; English.
XX The invention relates to a nucleic acid comprising a sequence at least 18
XX bases in length of a segment of the chemically pretreated DNA of genes
XX associated with pharmacogenetics according to one of the sequences of the
XX genes ALDH6 (NM 000693), CYP11A (NM 000781), CYP11B (NM 000497), CYP3A3
XX (NM 000776 and NM 017460), DPYD (NM 000110), EPHX2 (NM 001979), OCLN
XX (NM 002538), TXNRD1 (NM 003330), UGT8 (NM 003360), MRP (NM 004996,
XX NM 019900, NM 019901, NM 019902, NM 019862, NM 019898, NM 019899) and
XX their complementary sequences, or a sequence (S1) chosen from 87
XX sequences and their complements. The chemical pretreatment is bisulphite
XX treatment to convert cytosines (but not methyl-cytosines) into uracils.
XX Also included are an oligomer (II) in particular an oligonucleotide or a
XX peptide nucleic acid (PNA)-oligomer, comprising in each case at least one
XX base sequence having a length of 9 nucleotides which hybridises to or is
XX identical to a chemically pretreated DNA of genes associated with
XX pharmacogenetics and their complements, arranged in an array for
XX analysing diseases associated with the methylation state (CpG) and/or
XX detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The
XX oligomers may also be used as PCR primers. The set of 87 nucleic acids
XX and their complements is useful for diagnosis and therapy of solid
XX tumours and cancer. The present sequence represents one of the 87 DNA
XX sequences or its complement. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 6106 BP; 1938 A; 30 C; 849 G; 3289 T; 0 U; 0 Other;
Query Match 11.0%; Score 89.4; DB 6; Length 6106;
Best Local Similarity 51.7%; Pred. No. 0.0069;
Matches 278; Conservative 0; Mismatches 253; Indels 7; Gaps 3;
OY 198 TGTATGCGATTACTAATTTCTAGTATTAATGTTGGTTGGAAGTTAATATTAAT 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2612 TTTTATTTTAAATGTTTTGTTGTTTAAAGTTAAAGAAATTTAAATAGTTA 2671

OY 258 TCAATCCTGTTGTTTATTTTATTTTAAACAAATTAACAATTTATTTACTTG 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2672 TTAATGTTTAAATTTGTTTAAATTAATTAATTAATTAATTAATTAATGTTA 2731

OY 318 GTTGTTTCAATTTATGACGTTAATTTTATGTTTAAATAGACATTTATATAT 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2732 TATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2791

OY 378 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2792 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2850

```

QY 438 TATATTAATCAATCAAGATTAAGATTAATTAATCTTAA-----ATATTAAT 492
 Db 2851 TATTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2910
 QY 493 TTTTAAATAGCTTTTCAGTAAATCTGTCAACAAATGAATATTTTTCAGAGTTCAT 552
 Db 2911 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2970
 QY 553 CCAAAACCCGAAAGAAATGATTTTCAGAAAAGTAAATCATTTTCAGAAATTAATTT 612
 Db 2971 -TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3029
 QY 613 TCGAATATTTTTCAGGAAACAAATGAGCTTAAGTCTTCTGTTTATTTTAT 672
 Db 3030 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 3089
 QY 673 TTTTCTATTGAGAAACTGAAATGATTTGCAATGCTTAACTAGCTTGT 730
 Db 3090 TATAGTATTTTAATTAATTAATTTGCTTTTATGAGATTTATGTAATTTGTTT 3147

RESULT 13

ABL33472
 ID ABL33472 standard; DNA; 6106 BP.

XX ABL33472;
 AC
 DT 26-MAR-2002 (first entry)
 XX
 XX

Human immune system associated gene SEQ ID NO: 1445.

Human; immune system disease; cytosine methylation; antiaesthetic;
 antiarteriosclerotic; antianaemic; cytosolic; noctropic;
 neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 antihemematic; antiarthritic; antidiabetic; antipsoriatic;
 antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 ds.

XX Homo sapiens.
 OS
 XX
 XX

XX WO200200928-A2.
 PN
 XX
 XX

XX 03-JAN-2002.
 PD
 XX
 XX

XX 02-JUL-2001; 2001WO-EP007537.
 PF
 XX
 XX

XX 30-JUN-2000; 2000DE-01032529.
 PR
 XX
 XX

XX 01-SEP-2000; 2000DE-01043826.
 PR
 XX
 XX

XX (EPIC-) EPIGENOMICS AG.
 PA
 XX
 XX

XX Olek A, Piepenbrock C, Berlin K;
 PI
 XX
 XX

XX WPI; 2002-130909/17.
 DR
 XX
 XX

PT Nucleic acid comprising fragment of chemically modified gene, useful for
 diagnosis and treatment of diseases associated with abnormal cytosine
 methylation.
 PT
 XX

PS Claim 1; SEQ ID NO 1445; 32pp + Sequence listing; German.
 XX
 XX

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 CC
 XX

SQ Sequence 6106 BP; 1938 A; 30 C; 849 G; 3289 T; 0 U; 0 Other;
 Query Match 11.0%; Score 89.4; DB 6; Length 6106;
 Best Local Similarity 51.7%; Pred. No. 0.0069;
 Matches 278; Conservative 0; Mismatches 253; Indels 7; Gaps 3;

QY 198 TGTATGGGATTTCAATTTCTAGTATTAATTTGGTTTGAAGTAAATTAATTAAT 257
 Db 2612 TTTTATTTTAATAGTTTGTGTTTGAAGTTTGAAGAAATTTTAAATAGTTA 2671
 QY 258 TCAATCCTGTTTATTTTATTTTATTAACCAATTAATTAATTAATTAATTTG 317
 Db 2672 TTAATGTTTATTTAGTTTATTTTAAATTAATTAATTAATTAATTAATTAAT 2731
 QY 318 GTTGTTCATTTATGAGGTTAATTTTATTTTATTTTATTTTATTTATTTATTT 377
 Db 2732 TATATTAATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2791
 QY 378 AAATTAATTAATCAATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 437
 Db 2792 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2850
 QY 438 TATTAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 492
 Db 2851 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2910
 QY 493 TTTTAAATAGCTTTTCAGTAAATCTGTCAACAAATGAATATTTTTCAGAGTTCAT 552
 Db 2911 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2970
 QY 553 CCAAAACCCGAAAGAAATGATTTTCAGAAAAGTAAATCATTTTCAGAAATTAATTT 612
 Db 2971 -TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3029
 QY 613 TCGAATATTTTTCAGGAAACAAATGAGCTTAAGTCTTCTGTTTATTTTAT 672
 Db 3030 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 3089
 QY 673 TTTTCTATTGAGAAACTGAAATGATTTGCAATGCTTAACTAGCTTGT 730
 Db 3090 TATAGTATTTTAATTAATTAATTTGCTTTTATGAGATTTATGTAATTTGTTT 3147

RESULT 14

ADB54284
 ID ADB54284 standard; DNA; 6473 BP.

XX ADB54284;
 AC
 XX
 XX

XX 04-DEC-2003 (first entry)
 DT
 XX
 XX

XX Pretreated genomic DNA region 208.
 DE
 XX
 XX

XX colon cell proliferative disorder; non methylated CpG dinucleotide;
 XX cytosolic; cancer; adenoma; carcinoma; cytosine methylation state; ds.
 XX
 XX

XX Unidentified.
 OS
 XX
 XX

XX WO2003072821-A2.
 PN
 XX
 XX

XX 04-SEP-2003.
 PD
 XX
 XX

XX 27-FEB-2003; 2003WO-EP002035.
 PF
 XX
 XX

XX 27-FEB-2002; 2002EP-00004551.
 PR
 XX
 XX

XX (EPIC-) EPIGENOMICS AG.
 PA
 XX
 XX

XX Adorjan P, Burger M, Maier S, Nimrich I, Becker E, Lesche R;
 PI
 XX
 XX

XX WPI; 2003-731620/69.
 DR
 XX
 XX

PT Detecting and differentiating between colon cell proliferative disorders
 PT associated with a gene or its regulatory regions comprises contacting a
 PT target nucleic acid in a biological sample obtained from the subject with
 PT a reagent.

XX Claim 32; SEQ ID NO 340; 74bp; English.

XX The invention relates to a novel method for detecting and differentiating
 CC between colon cell proliferative disorders associated with at least one
 CC gene or its regulatory regions. The method comprises contacting a target
 CC nucleic acid in a biological sample obtained from the subject with at
 CC least one reagent or a series of reagents, where the reagent or series of
 CC reagents, distinguishes between methylated and non methylated CpG
 CC dinucleotides within the target nucleic acid. The molecules of the
 CC invention demonstrate cytosatic activity whilst the method may useful
 CC for detecting and differentiating between colon cell proliferative
 CC disorders, including cancers such as colon adenoma and colon carcinoma.
 CC The PNA (peptide nucleic acid)-oligomers are useful as probes for
 CC determining cytosine methylation state or single nucleotide
 CC polymorphisms. The current sequence is that of the pretreated genomic DNA
 CC region of the invention. This sequence is not shown within the
 CC specification but is taken from Wipoweb.

XX Sequence 6478 BP; 1878 A; 0 C; 1282 G; 3318 T; 0 U; 0 Other;

XX Query Match 10.9%; Score 89; DB 9; Length 6478;

XX Best Local Similarity 51.5%; Pred. No. 0.0076; Mismatches 235; Indels 2; Gaps 2;

XX Matches 252; Conservative 0; Mismatches 235; Indels 2; Gaps 2;

XX 220 TAGTATTAAATGTTGGTGGAGTTAAATATATATATATCAATCCCTGTTTATTTT 279

XX 4349 TAGTATGGTGTATTAATATATATATATATATATATATATATATATATATATAT 4408

XX 280 TTTTATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 339

XX 4409 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4467

XX 340 TATATTTAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 399

XX 4468 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4527

XX 400 TATGTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 459

XX 4528 TTTGAAATTTGATTTTGTATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTA 4587

XX 460 TAAAAGTAAATTAATTTCTTAATTAATTAATTTTAAATTTTAAATTTTAAATTTG 519

XX 4588 TAAATTAATTAATTTTAAATTAATTAATTTGTAATTAATTAATTAATTAATTAATTA 4647

XX 520 TCAACATTAATTAATTTTTCAGGTTCAATCCAAACACCAAGAAATTAATTAATTTT 579

XX 4648 AGAAAGTAAATTAATTAATTTTGTATTAATTAATTAATTAATTAATTAATTTT 4707

XX 580 -CAGAAAGTAAATTAATTTTCAGAAATTAATTTTCGAAATTAATTTTACTGGCAACA 638

XX 4708 AATTAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 4767

XX 639 ATGAGTCTAAGTCTTCTGTTTATTTTATTTTCTATTGAGAAATTAATTTT 698

XX 4768 TAAATTTTGAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTA 4827

XX 699 GATTGTCA 707

XX 4828 TATTTTAA 4836

XX RESULT 15

XX ABL32727 standard; DNA; 16602 BP.

XX ABL32727;

XX 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 700.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antiasthmatic; cytosatic; necrotic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antirheumatic; antidiabetic; antipsoriatic;
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 KW ds.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JAN-2002.

XX 02-JUL-2001; 2001WO-EP007537.

XX 30-JUN-2000; 2000DE-01032529.

XX 01-SEP-2000; 2000DE-01043826.

XX (EPig-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.

XX Claim 1; SEQ ID NO 700; 32bp + Sequence listing; German.

XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention

XX Sequence 16602 BP; 4432 A; 202 C; 3755 G; 8213 T; 0 U; 0 Other;

XX Query Match 10.8%; Score 88.2; DB 6; Length 16602;

XX Best Local Similarity 48.8%; Pred. No. 0.008; Mismatches 308; Indels 8; Gaps 2;

XX Matches 301; Conservative 0; Mismatches 308; Indels 8; Gaps 2;

XX 124 GCATTAACTGTTTGGTGAATTTAATGAGTAAATTTCAATTAATTTCAATGAAG 183

XX 15284 GTATTAAATTAATTTAATGTTTGTAGTAGTGTGGAGATTTTATAGAGATGTT 15343

XX 184 GTTTAAGATTAATTAATTTGATTAATTTAATTTAATTTAATTTAATTTAATTTAAT 243

XX 15344 TTTAATTAATTAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 15403

XX 244 TAAATTAATTAATTTCAATTTCTGTTTAAATTTTAAATTTTAAATTTAATTTAAT 303

XX 15404 TCGTGTGGAAGTTTAAATTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTAT 15463

XX 304 TTTAATTAATTTGTTGTTTCAATTTATGCGGTTAAATTTAATTTAATTTAATTTAAT 363

XX 15464 TTTTATTTTATTTGTTAATTAATGAGTTAGTTATTTGTTTATTTAATTTAGTTGGAG 15523

XX 364 CATTAATTAATTAATTAATTAATTAATTTCAATTAATTAATTAATTAATTAATTA 423

XX 15524 GTGTTTAAATTTTAAAGAAATTTGTAATTAATTTAATTAATTTAATTTAATTTA 15583

XX 424 AATTAATTAATTAATTAATTAATTTCAATTAATTAATTAATTAATTAATTTAATTT 483

XX 15584 GATTATTTTAAATTTGTTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAAT 15643

```
QY 484 ATATAATTTTAAATAGCTTTCAGTAATCTGCAACAATAGAAAAATTTTGG 543
Db 15644 AATATAGT-----AATAGATAGAAAAATATAATATATATATATGTAATTTATT 15696
QY 544 CAGGTCATCCAAACACACAGAAAGTAATCATTTTCAGAAAGTAATCATTTTCAGA 603
Db 15697 TATGTTAATTTTATAGAGAGATTGAGATTATTTTGTATAGTATATATTTTATTAT 15756
QY 604 AATTATTTTTCGAAATTATTTTACTGCAACAAT -GGAGTCTAGTGTCTGTTT 662
Db 15757 TTTTATGTTTTTTAAATTTTATTTTATTTTATATAGTTTATAGTTTATTTTGCTT 15816
QY 663 TATTTTATTTTCTATTTAGAGAACTAGAAATGATTTGCAAAATGCTTTAATCTA 722
Db 15817 TATTTTGGTTTATTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTATTTATTTT 15876
QY 723 GCTTGTATGATTAGTT 739
Db 15877 TTTTATTTAGTTTTT 15893
```

Search completed: September 20, 2004, 20:51:25
Job time : 411 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comogen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2004, 20:44:50 ; Search time 81 Seconds
(without alignments)
5590.616 Million cell updates/sec

Title: US-10-089-557-2_COPY_1_816

Perfect score: 816
1 actatagggcagcgtggtc.....tatttcagttgtaaaag 816

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88.2	10.8	5852	1	US-07-867-106-2
2	82.4	10.1	665	2	US-08-883-795A-36
3	78.4	9.6	6124	4	US-08-213-419B-3
4	78	9.6	5852	1	US-07-867-106-2
5	77	9.4	5152	4	US-10-204-708-74
6	76.2	9.3	665	2	US-08-883-795A-36
7	75.8	9.3	5455	4	US-10-204-708-33
8	75.6	9.3	6156	4	US-10-204-708-60
9	75.2	9.2	20674	4	US-09-641-638-651
10	74.8	9.2	5152	4	US-10-204-708-73
11	74.4	9.1	6669	4	US-10-204-708-6
12	74.2	9.1	2058	2	US-08-749-391-1
13	74.2	9.1	2058	3	US-09-390-200-1
14	73.8	9.0	20674	4	US-09-641-638-651
15	73.6	9.0	6070	4	US-10-204-708-10
16	73.2	9.0	19124	2	US-08-487-826B-13
17	73.2	9.0	640681	4	US-09-790-988-1
18	73	8.9	11049	4	US-10-204-708-22
19	72.2	8.8	3701	3	US-08-845-258-1
20	72.2	8.8	3701	3	US-08-990-571-10
21	72.2	8.8	3701	4	US-08-723-142A-10
22	72.2	8.8	3701	4	US-09-528-784A-10
23	72.2	8.8	3701	4	US-09-569-098A-10
24	72.2	8.8	6243	4	US-09-056-075-1
25	72	8.8	53332	4	US-09-801-861-3
26	71.8	8.8	834	3	US-08-998-416-305
27	71.8	8.8	5152	4	US-10-204-708-73

28	71.6	8.8	7832	4	US-09-004-838-94	Sequence 94, Appl
29	71.2	8.7	6317	4	US-10-204-708-11	Sequence 11, Appl
30	71.2	8.7	26000	4	US-09-843-376-10	Sequence 10, Appl
31	70.8	8.7	10640	4	US-09-417-485D-5	Sequence 5, Appl
32	70.6	8.7	5152	4	US-10-204-708-74	Sequence 74, Appl
33	70.6	8.7	6317	4	US-10-204-708-11	Sequence 11, Appl
34	70.4	8.6	2435	3	US-09-306-593-1	Sequence 2, Appl
35	70.4	8.6	10467	4	US-10-204-708-2	Sequence 1000, Ap
36	70.2	8.6	640681	4	US-09-790-988-1	Sequence 186, App
37	70.2	8.6	1441	4	US-08-956-171E-1000	Sequence 288, App
38	69.6	8.5	617	3	US-08-998-416-186	Sequence 8, Appl
39	69.6	8.5	835	3	US-08-998-416-288	Sequence 8, Appl
40	69.6	8.5	1511	1	US-07-991-867B-8	Sequence 8, Appl
41	69.6	8.5	1511	1	US-08-107-755A-8	Sequence 8, Appl
42	69.6	8.5	1511	2	US-08-544-332-8	Sequence 8, Appl
43	69.6	8.5	1511	4	US-09-370-861A-8	Sequence 23, Appl
44	69.4	8.5	11049	4	US-10-204-708-23	Sequence 13, Appl
45	69	8.5	642	1	US-08-764-100-13	

ALIGNMENTS

RESULT 1
US-07-867-106-2
Sequence 2, Application US/07867106

Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526rits
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feehey, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS

LOCATION: 2378..5038
US-07-867-106-2

Query Match 10.8%; Score 88.2; DB 1; Length 5852;
Best Local Similarity 48.9%; Pred. No. 2e-07;
Matches 297; Conservative 0; Mismatches 303; Indels 7; Gaps 2;

QY 147 TTAAAGGAGGCTAATGCGAAGATTTTCATGGAAGGTGTAAGATTAATGTTATG 206
DB 1716 TTGATGACATTTATTTCCAGTTTCTTTTAAATGAGATAAATAATGCTATC 1775
QY 207 GATTACTAATTTCTAGATTAATTTGGTTGGAAGTAAATATTAATTAATTCAC 266
DB 1776 GATATATACTTAAT-TTATTAAGATTAATTAATTTTAAATTAATTCACCTT 1834
QY 267 GTTTTATTTTCTTTTAAACAAATTAACAAATTAATTAATTAATTTGTTTTC 326
DB 1835 TTTTCTTTTCTTTTAAATTTTAAATTTTCTTTTAAATTTTCTTTTAAATTT 1894
QY 327 AATTATGACGTTAATTTTGTATTAATTAATTAATTAATTAATTAATTAAT 386
DB 1895 TTAAATATCAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1954
QY 387 AATCATTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 446
DB 1955 TGTAGTTTAAACACTTTTCTATTTCTTTTCTTTTCTTTTCTTTTCTTTGA 2014
QY 447 CTCATCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 506
DB 2015 AATTAAT 2074
QY 507 TTCAGTAATCTGTCAACATTAATTAATTTT-----TTCAGTATCAACAC 560
DB 2075 TTTTGTATTTTCTTTTAAATTTTAAATTTTCTTTTCTTTTCTTTTCTTTTCT 2134
QY 561 CAGAAAGTAATCAATTTTCAAGAAATTAATTTTCAAGAAATTAATTTTCTTGA 620
DB 2135 ATAAATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2194
QY 621 TATTTTACGCAACAAATGAGTCTAGTCTGTTTCTTTTCTTTTCTTTTCTTAT 680
DB 2195 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2254
QY 681 TTAGAGAACTGGAATGATTTGCAATGCTTAATCTGCTTTTAAATTAATTTG 740
DB 2285 TTTTAAATTTTAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2314
QY 741 AAGGCA 747
DB 2315 AATTTCA 2321

RESULT 2

US-08-983-795A-36/c
Sequence 36, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcive, Genevieve
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
NUMBER OF INVENTION: Vectors for Tissue Plasminogen Activator
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline

REGISTRATION NUMBER: 40,261

TELEPHONE: (416) 364-7311

TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 665 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

CLONE: Rh 32

US-08-883-795A-36

Query Match 10.1%; Score 82.4; DB 2; Length 665;
Best Local Similarity 49.4%; Pred. No. 1.7e-06;
Matches 244; Conservative 0; Mismatches 246; Indels 4; Gaps 1;

QY 242 GTTAATATTAATTAATTAATTAATTAATTTTCTTTTAAATTAATTAATTAAT 301
DB 599 GTCTTTTAAATTTTCTGCTTTTAAATTAATTAATTAATTAATTAATTAAT 540
QY 302 AATTTATTAATTTGTTGTTTCAATTAATTAATTAATTAATTAATTAATTAAT 361
DB 599 GTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 362 AGCATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 421
DB 479 GTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
QY 422 TAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 481
DB 419 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360
QY 482 AATATTAATTT-----TTTAAATTAATTTTCTTTTCTTTTCTTTTCTTTTCT 537
DB 359 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
QY 538 TTTTGCAGTTCAATCAACACAGAAAGTAATCAATTTTCAAGAAAGTAATCAATTT 597
DB 299 TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
QY 598 TTCAAAATTAATTTTGGAAATTAATTTTCTGCAACAAATGAGTCTAAGTCTTCT 657
DB 239 TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
QY 658 GTTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 717
DB 179 TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120
QY 718 ATCTAGCTGTTTA 731
DB 119 TTATTAATTAATTTA 106

RESULT 3

US-08-213-419B-3
Sequence 3, Application US/08213419B
Patent No. 6333406
GENERAL INFORMATION:
APPLICANT: Inselburg, J. et al.
TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM

TITLE OF INVENTION: Improved Plasmid Vectors for Cellular

QY 566 A8J1AALC1111C6MNNOLNRR.CAR.....

Db 5490 CAATAAACATATTGATTTTTTTTTTTTTTTTTTTTTTTTAAAAATTCAAATTAAT 54311
 QY 626 TACTGGCAACAAGATCGAGCTAGATGTTCTGTTTTAT--TTTATATTTTCATATTA 683
 Db 5430 AAAATATATTCTATATATCTGGAAGAACTTCAATTTTATTAATTAATATATTAATTT 5371
 QY 684 GAGAACTAGAAATGATTGTGCAATGCTTTATCTAGCTGTGATAGTAACTAGAG 743
 Db 5370 TAAAAATTAGATCTATCGATCTATATTAATTTCCATGTTTTTAAATTTTTTTTAAAG 5311

RESULT 5
US-10-204-708-74/

```

sequence 1/4, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
TITLE OF INVENTION: by Assessing DNA Methylation
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 74
LENGTH: 5152
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-74

```

Query Match	9.48;	Score 77;	DB 4;	Length 5152;
Best Local Similarity	49.7%;	Pred. No. 1.8e-05;		
Matches 252;	Conservative 0;	Mismatches 250;	Indels 5;	Gaps 2;
QY	225	TTAATGTGGTTGGAGTTAATATATATATATGCAATCGTTGTTTTATTTTTTTTT	284	
Db	2165	TTAAATACCAATTCACACAAATCCTTTTAAATCTAAACAAACACTTTTACCATATATA	2106	
QY	285	ATAACACAATTACAAATATATTATTTAACTTGTTGGTTTCATTTATGACGGTTATA	344	
Db	2105	ATTATATATATATATATATTTTAAACTTATTTTACTTTATATATATATTAATTAATTT	2046	
QY	345	TTTTGCTTAAATATGAGCAATTATATATATTAATTAATTAATTAATCATGTAATATATGT	404	
Db	2045	TTTTAAATACCTACGTAAATTTTAAACATTAACATTAACATTAACATTAATTAATTAATTA	1986	
QY	405	AAAAATATTT--TAAATATTAATTTATATATATATATTAATTAATTAATCAATCAACAATA	461	
Db	1985	CATATCCCTTCCTCAAAAAAATTCAAATATACCTATTCAACACAAACAAACAAATTA	1926	
QY	462	AAAAGATATTAATTTCTTAAATATATATATTTTTTAAATAGCTTTTCAGTAATCTGTC	521	
Db	1925	ATTAATACCTTAACGTATTAATAAAAAATTTTAAAAAACCAATATTATAAAAATATATA	1866	
QY	522	AAACAAATAGAAAAATTTTTTTGGACGGTCAATCCAAACACAGAAAGTAAATCATCTTTCA	581	
Db	1865	AATATATCTTAATATATTTTTTACAAATATATTTTTTAAATCACTAAATTAATCTTATTTTA	1806	
QY	582	GAATAGTAA--ATCATTTTTCAGAATTAATTTTTTCGAAATATTTTACTCGCAACAA	639	

Db 1805 AATTAACCACTATTAACATTTTAAAAATTTTATTAACAATTACTTTCCAAATTAATCTC 1746

Qy 640 TGGAGTCAGAGTGTCTCGTTTATTTTATTTTCTATTAGAGAACTAGAAATTG 699

Db 1745 GTTATATCAAAATACTTATCTTCATCGATCATTTTATATCTAAAAAATTTTTAAAAA 1686

Qy 700 ATTTGTCAAATGCTTTAATCTAGCTT 726

Db 1685 AAAAAACAATTAATTAATAATTTT 1659

RESULT 6

```

US-08-883-795A-36
Sequence 36, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcove, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasmidogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PAR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Michelle
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO.: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: Rh 32
US-08-883-795A-36

```

Query Match	9.3%	Score 76.2	DB 2	Length 665
Best Local Similarity	47.1%	Pred. No. 2.1e-05		
Matches 268	Conservative 0	Mismatches 298	Indels 3	Gaps 1
QY	184	GTGTTAAGATTAAATGATTATGGGATTAAGTCTATTTTCAGTATTAAATGTTGGTTGGAACT	243	
DB	20	GTATTTTAAATTAATTAATTTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	79	
QY	244	TAAATATTAATTAATTCACCTCGTGTATTTTATTTTATTTTATTAACAATTAACAATTA	303	
DB	80	TATTTTAAATTAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	139	
QY	304	TTTATTTAACTTGGTGTGTTTCATTTATGACGGTAAATTTTGTTTAATTAATG	363	
DB	140	TATTTTAAATTAATTAATTTATTAATTAATTTTATTAATTAATTAATTAATTTTAAAT---	196	

Oy		364	CATTATTAATATAATAAATAAATCACTGGTATATATGTAAAAATATTTAAAATATA	423
Db		197	AAATATTTTAAATTAATAATTTATTAATTAATTAATTTTAAATTAATTAATTTAAT	256
Oy		424	AATTTAATTAATATAATATAACTCAATCAAACAATAAAAAGATAATTAATTCGTAAT	483
Db		257	AAATATTTTAAATTAATAATGTTTTAATTAATTAATTTTATATTAATAATGTTTAAAT	316
Oy		484	ATATPAATTTTTTAAATAGCTTTCCAGTAATTCGTCAACATAGAATAATTTTTTG	543
Db		317	ACATATTTTAAATTAATAATGTTTTTAATTAATTAATTTTATATAATTAATAATGTTTAAT	376
Oy		544	CAGGTTCATCCAAACACCAGAAAGTAATCATTTTCAGAAAAGTAATCATTTTTGCAG	603
Db		377	ACATATTTTAAATTAATAATGTTTAAATTAATTAATTAATTAATTAATAANGTTTAAAT	436
Oy		604	AATTAATTTTTGGGAATATTTTACTCGCAACAAATGAGTCTAAGTCTGTTTT	663
Db		437	ACATATTTTAAATTAATTAATTAATTTTAAAGTATTTTAAATTAATTAATTAATTAATAAG	496
Oy		664	ATTTTATATTTTCTATTTTAGAGAAATCAGAAATTAATTTGTCAAATGCTTAATCTAG	723
Db		497	TATTTATATTAATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT	556
Oy		724	CTTGTTAGATTAGTTGAAGGCCAGAA	752
Db		557	ATTTTATTAATTAATTAATAAGACGAGAA	585

RESULT 7
US-10-204-708-33
: Sequence 33, Application US/10204708

```

1  APPLICANT: BERLIN, KURT
2  TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
3  TITLE OF INVENTION: by Assessing DNA Methylation
4  FILE REFERENCE: 503.3.1012
5  CURRENT APPLICATION NUMBER: US/10/204,708
6  CURRENT FILING DATE: 2003-05-06
7  PRIOR APPLICATION NUMBER: PCT/EP01/03971
8  PRIOR FILING DATE: 2001-04-06
9  PRIOR APPLICATION NUMBER: DE 10019058.8
10 PRIOR FILING DATE: 2000-04-06
11 PRIOR APPLICATION NUMBER: DE 10019173.8
12 PRIOR FILING DATE: 2000-04-07
13 PRIOR APPLICATION NUMBER: DE 10032529.7
14 PRIOR FILING DATE: 2000-06-30
15 PRIOR APPLICATION NUMBER: DE 10043826.1
16 PRIOR FILING DATE: 2000-09-01
17 NUMBER OF SEQ ID NOS: 98

```

```

? SEQ ID NO 33
?
? LENGTH: 5455
?
? TYPE: DNA
?
? ORGANISM: Artificial Sequence
?
? FEATURE:
?
? OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
?
? US-10-204-708-33

```

Query Match	9.3%	Score 75.8;	DB 4;	Length 5455;
Best Local Similarity	46.7%	Pred. No. 2.9e-05;		
Matches 274;	Conservative	0;	Mismatches 312;	Indels 1;
				Gaps 1

QY 160 TATGCGAGTATTTTTCATGCAAGCGTTAGAAATTAATGTTATGGAGTTCTCAATTTTC 219

Db 568 TTTTATTTTATTTATTTTGGTAGTACGTAAATTAACGTTTTTTTGGTATTTTATTTT 627

QY 220 TAGATTAATGTGCTTTGGAGATTAAATTAATTAATTTTCAATCCGTGTTTATTTT 279

Db 628 TTGTTTTTTTATTTAGTTTGGGAGATTAAATGATTATTCGATTAATTTTATAGTTTTC 687

QY	280	TTTTTATACAACTTACAAATAATTTATTTAACTTGGTGGTTTCAATTTATGACGGT	339
Db	688	TTTTTAAATTTGCAATTTAGTTTTGTTTTGTTTGGTTTTTTTAAAGTTTTTAATTTA	747
QY	340	TAATATTTTA-GTTTAAATTAATTGACATTAATTAATATTAATTAATAATCAATGTAAT	398
Db	748	TTTGTTTTTTATTTTGTAGCTTTTTTTTAAATATCGTTTTTTTTTGAAGCGTTTTTAATTA	807
QY	399	ATATGTAAAAATTAATTTAAAAATATAATTAATTAATTAATTAATAATCAATCAACACA	458
Db	808	ATAATATATATAATATAATATAATATAATATAATATAATTTTTTTATGTATATTTTA	867
QY	459	ATAAAAGATATAATTTCTTAATATAATATAATTTTTTAAATAGCTTTTCAGTAATCT	518
Db	868	TTTTTACATTTTGTAGGGAATATAGTTTAAATGAGTGTTTTTTTTGTTTTTTTAAAT	927
QY	519	GTCAAAACATAGAAATAATTTTTTTCAGGCTCATCCAAACCGAAGAAAGTAATCAATT	578
Db	928	TTGTGTGGTTTTTATGTTTTTTTTTTATTAATGTCGTTTTTTTAATTTATTTAGTTTAA	987
QY	579	TCAGAAAGTAATCAATTTTCAGAAATTAATTTTTTCGAAATTAATTTTACTCGCAAAACA	638
Db	988	TTGGGAATTTTTTTTTTTTTTTTATTTTATTTTTTTTTTGTAGAAATTTGTAATTTTT	1047
QY	639	ATGAGCTTAAGTGTTCCTGTTTTTATTTTTTATTTTTTCAATTTAGAGAAACTGAAAT	698
Db	1048	ATAAAATTTTTTGAAGGTGATTTTGAATTTTAATTTTGTATTAATGTTAATGTTTTTT	1107
QY	699	GATTTGTCAAAATGCTTTAATCTAGCTGTTTGAATTAAGTGMAGGG	745
Db	1108	TTTTTTAAATTAATTTTTTAAATTAAGAGTGTACGTGGGATTTGGGGG	1154

Db 1108 TTTTAAATTTATTTTAACTAAGAGTGACGTGGGATTTGGGG 1154

```

RESULT 8
US-10-204-708-60
Sequence 60, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIPEPENROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
TITLE OF INVENTION: by Assessing DNA Methylation
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 60
LENGTH: 6156
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-60

```

Query Match	9.3%	Score 75.6;	DB 4;	Length 6156;
Best Local Similarity	49.0%	Pred. No. 3.2e-05;		
Matches 291; Conservative	0;	Mismatches 294;	Indels 9;	Gaps 3;

QY 155 ACGGCTATGTCCAGATATTTCATGGAAAGTGTAAGCAATTATGTTATGGATACIA 214

Db 5222 ATTGTATTGTTGTTTATTATAGTGTTATTATAATATTAATTGTTAGATTITGGTA 5281

```

1 NAME/KEY: exon
2 LOCATION: 3871..4072
3 OTHER INFORMATION: exon 2
4 NAME/KEY: exon
5 LOCATION: 5552..5633
6 OTHER INFORMATION: exon 3
7 NAME/KEY: exon
8 LOCATION: 5758..5880
9 OTHER INFORMATION: exon 4
10 NAME/KEY: exon
11 LOCATION: 5996..6099
12 OTHER INFORMATION: exon 5
13 NAME/KEY: exon
14 LOCATION: 6349..6509
15 OTHER INFORMATION: exon 6
16 NAME/KEY: exon
17 LOCATION: 7379..7522
18 OTHER INFORMATION: exon 7
19 NAME/KEY: exon
20 LOCATION: 8645..8854
21 OTHER INFORMATION: exon 8
22 NAME/KEY: exon
23 LOCATION: 12254..12340
24 OTHER INFORMATION: exon 9
25 NAME/KEY: exon
26 LOCATION: 12854..13023
27 OTHER INFORMATION: exon 10
28 NAME/KEY: exon
29 LOCATION: 13308..13429
30 OTHER INFORMATION: exon 11
31 NAME/KEY: exon
32 LOCATION: 16567..16667
33 OTHER INFORMATION: exon 12
34 NAME/KEY: exon
35 LOCATION: 16775..16945

```

```

NAME/KEY: exon
LOCATION: 3871..4072
OTHER INFORMATION: exon 2
NAME/KEY: exon
LOCATION: 5552..5633
OTHER INFORMATION: exon 3
NAME/KEY: exon
LOCATION: 5758..5880
OTHER INFORMATION: exon 4
NAME/KEY: exon
LOCATION: 5996..6099
OTHER INFORMATION: exon 5
NAME/KEY: exon
LOCATION: 6349..6509
OTHER INFORMATION: exon 6
NAME/KEY: exon
LOCATION: 7379..7522
OTHER INFORMATION: exon 7
NAME/KEY: exon
LOCATION: 8645..8854
OTHER INFORMATION: exon 8
NAME/KEY: exon
LOCATION: 12254..12340
OTHER INFORMATION: exon 9
NAME/KEY: exon
LOCATION: 12854..13023
OTHER INFORMATION: exon 10
NAME/KEY: exon
LOCATION: 13308..13429
OTHER INFORMATION: exon 11
NAME/KEY: exon
LOCATION: 16567..16667
OTHER INFORMATION: exon 12
NAME/KEY: exon
LOCATION: 16775..16945
OTHER INFORMATION: exon 13
NAME/KEY: exon
LOCATION: 17063..17554
OTHER INFORMATION: exon 14
NAME/KEY: misc feature
LOCATION: 17555..20674
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 1128
OTHER INFORMATION: 10-508-191 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1182
OTHER INFORMATION: 10-508-245 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1559
OTHER INFORMATION: 10-509-284 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1570
OTHER INFORMATION: 10-509-295 : deletion of C
NAME/KEY: allele
LOCATION: 1827
OTHER INFORMATION: 10-510-173 : variable motif ATTGA or TTTT
NAME/KEY: allele
LOCATION: 2048
OTHER INFORMATION: 10-511-62 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2323
OTHER INFORMATION: 10-511-337 : insertion of T
NAME/KEY: allele
LOCATION: 2341
OTHER INFORMATION: 10-512-36 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 2623
OTHER INFORMATION: 10-512-318 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2832
OTHER INFORMATION: 10-513-250 : polymorphic base A or G
NAME/KEY: allele

```

LOCATION: 2844 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2934 : polymorphic base A or G
OTHER INFORMATION: 10-513-352 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2947 : polymorphic base A or G
OTHER INFORMATION: 10-513-365 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 3802 : polymorphic base A or G
OTHER INFORMATION: 12-206-81 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 4062 : deletion of C
OTHER INFORMATION: 10-343-231 : deletion of C
NAME/KEY: allele
LOCATION: 4088 : polymorphic base C or T
OTHER INFORMATION: 12-206-366 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4109 : polymorphic base C or T
OTHER INFORMATION: 10-343-278 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4170 : polymorphic base G or T
OTHER INFORMATION: 10-343-339 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 5903 : polymorphic base A or G
OTHER INFORMATION: 10-346-23 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6019 : polymorphic base A or G
OTHER INFORMATION: 10-346-141 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6141 : polymorphic base G or C
OTHER INFORMATION: 10-346-263 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6183 : polymorphic base C or T
OTHER INFORMATION: 10-346-305 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6338 : polymorphic base A or G
OTHER INFORMATION: 10-347-74 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6375 : polymorphic base G or C
OTHER INFORMATION: 10-347-111 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6429 : polymorphic base C or T
OTHER INFORMATION: 10-347-165 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6467 : polymorphic base A or G
OTHER INFORMATION: 10-347-203 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6484 : polymorphic base A or G
OTHER INFORMATION: 10-347-220 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6534 : polymorphic base A or T
OTHER INFORMATION: 10-347-271 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 6611 : polymorphic base A or G
OTHER INFORMATION: 10-347-348 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 7668 : polymorphic base A or G
OTHER INFORMATION: 10-348-391 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8608 : polymorphic base C or T
OTHER INFORMATION: 10-349-47 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 8658 : polymorphic base A or G
OTHER INFORMATION: 10-349-97 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8703 : polymorphic base G or C
OTHER INFORMATION: 10-349-142 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 8777 : deletion of CTG
OTHER INFORMATION: 10-349-216 : deletion of CTG
NAME/KEY: allele
LOCATION: 8785

OTHER INFORMATION: 10-349-224 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 8926 : polymorphic base C or T
OTHER INFORMATION: 10-349-368 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12171 : polymorphic base C or T
OTHER INFORMATION: 10-350-72 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12429 : polymorphic base C or T
OTHER INFORMATION: 10-350-332 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13341 : polymorphic base A or G
OTHER INFORMATION: 10-507-170 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 13492 : polymorphic base A or C
OTHER INFORMATION: 10-507-321 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 13524 : polymorphic base C or T
OTHER INFORMATION: 10-507-353 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13535

Query Match 9.2%; Score 75.2; DB 4; Length 20674;
Best Local Similarity 50.4%; Pred. No. 4.2e-05;
Matches 266; Conservative 0; Mismatches 253; Indels 9; Gaps 3;

QY 162 TGTCGAAGTATTTTCATGAGGCTTAAAGATTATGTTGGAGTTACTAATTTCTA 221
11581 TATTTAAAGGAGCCATGCTTTTATATTAATAATTTTCTTACCTATTAATAAT 11522

QY 222 GTATTAATGCTGTTTGAGAGTAAATATATATATTCATCCCTGTTTATTTT 281
11521 TTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11462

QY 282 TTATTAACACATTAACAAATTAATTTTAACTTGTGCTTTTCAATTATGAGGTTA 341
11461 TTATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 11407

QY 342 ATATTTAGTTTAATAATAATGAGCATTAATTAATTAATTAATTAATTAATTAAT 401
11406 AAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11347

QY 402 TGTAAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 461
11346 TTAATAATTAATTAATAAT -TAAATTAATTAATAATAATAATAATAATAATAAT 11288

QY 462 AAAGATTAATAATAATCTTAATTAATAATAATAATAATAATAATAATAATAATA 521
11287 TAACTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 11228

QY 522 AAACATTAATAATAATTTTTCAGAGGTTCAATCCACACAGAAAGTAATCAATTTCA 581
11227 ATTTAAATAATAATAATTAATAATAATAATAATAATAATAATAATAATAATAAT 11171

QY 582 GAAAGTAATCAATTTTTCAGAAATTAATTTTTCGGAATTAATTTTACGCGAAACAATG 641
11170 TAAATTAATAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAAT 11111

QY 642 GAGCTAAGTGTCTGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 689
11110 TTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 11063

RESULT 10
US-10-204-708-73
Sequence 73, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
TITLE OF INVENTION: by Assessing DNA Methylation

```
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 73
LENGTH: 5152
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-73
```

```
Query Match
Best Local Similarity 49.6%; Score 74.8; DB 4; Length 5152;
Matches 277; Conservative 0; Mismatches 272; Indels 9; Gaps 3;
```

```
173 TTTCATGAGGCTGTAGAAATTAATGTTGGGATTACTAATTTCTAGATTAATTTGT 232
2942 TTTTAAATATATTTTGAATATGATTTAGTTGTATTTATTTATTTAGTTT 3001
233 GGTTCGAGTTAATATATATTTCAATCTGTTGTTTATT--TTTTTTTATAA 290
3002 TATATTAATTTATTTTAAATTTGTAATTTTATTTATTTAGTTGTTTATATA 3061
291 CAATTACAAATATTTATTTACTTTGTTGTTTCAATTTTGAAGTTAATTTTG 350
3062 TATGTTTAAATGATTTATTTTATTTATATGTTGATTTGATTAATTTTGGATTT 3121
351 TTTAATTAATGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 410
3122 TATGATTTAGTATATATTTAGTATTTAGGTTTAAATTTGGG---ATATATTTT 3176
411 AATTTAAATTAATTTTATATATTTATTAATTAATTAATTAATTAATTAATTA 470
3177 TTTTGTAGAAATTTTATATGTTTGTAGATTAATTAATTAATTAATTAATTA 3236
471 TAAATCTTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 530
3237 TAAATTTGAGATTAATTAATTTGAGATTTGATTTGAGATTTAATTAATTTG 3296
531 AAAATATTTTTCAGAGTTCAATCAAAACAGAAAGTAATCAATTTTCAGAAA 590
3297 GAATGTTTATTAATTAATTTTAAAGTTATGATGATTTAGTTTGAATTAAG 3356
591 --ATCATTTTTCAGAAATTTTTCGGAATTTTCTGCGCAAAACAATGAGCTCA 648
3357 TTGTTAATTAATTTTAAATTAATTTTGTGTAATTTTAAATTTTAAATTTT 3416
649 AGTGTTCGTTTATTTTATTTTATTTTCTATTTAGAGAACTAGAAATTTGATG 708
3417 AATATGTTTATTTTACGATTTATTTAGTATTTAGAAAAGAAAAGAAAAGTAA 3476
709 ATGCTTTAATCTAGCTT 726
3477 GTAATTAAGGTGATTT 3494
```

```
RESULT 11
US-10-204-708-6
; Sequence 6, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
```

```
APPLICANT: PIEDEBERCK, Christian
TITLE OF INVENTION: Berlin, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
TITLE OF INVENTION: by Assessing DNA Methylation
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 6
LENGTH: 6669
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-6
```

```
Query Match
Best Local Similarity 48.4%; Score 74.4; DB 4; Length 6669;
Matches 237; Conservative 0; Mismatches 251; Indels 2; Gaps 1;
```

```
184 GTGTTAAGATTAATGATTTATTTGGGATTAATTTCTAGATTAATTTGTTGGAGT 243
6057 GTATTTATTTGATTTTAAATTAATTAATTAATTTTAAATTTTAAATTTTAA 6116
244 TAAATTAATTAATTTCAATTTCTGTTTATTTTATTTTATTTTATTTTATTTA 303
6117 TTTTATTTTATTAAGAGATTTATTTATTTTATTTTATTTTATTTTATTTTAT 6176
304 TTTATTAATTTGTTGTTTCAATTTATG--ACGGTTAATTTTATTTTATTTATTT 361
6177 GTATTTTGTATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 6236
362 AGCATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 421
6237 GTTTTGTATTTTGTGATTTGATTAATTAATTAATTAATTAATTAATTAAT 6296
422 TAAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 481
6297 TATTTGATTTTAAAGATTAATTAATTAATTTTATTTTATTTTATTTTATTT 6356
482 ATATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 541
6357 AAGGAGATTTTATTAATTAATTAATTAATTTTATTTTATTTTATTTTATTT 6416
542 TCGAGTTTCAATCAACCCAGAAAGTAATCAATTTTCAGAAAAGTAATCAATTT 601
6417 TATTTTATTTTATTAATTAATTAATTAATTAATTTTATTTTATTTTATTT 6476
602 GAATTTATTTTTCGAAATTTATTTTACGCGCAAAACAATGAGCTGATTTGCTG 661
6477 TTTTATTTTATTTTTCGTTGTTTATTTTATTTTATTTTATTTTATTTTATTT 6536
662 TTAATTTTAA 671
6537 ATTTATATA 6546
```

```
RESULT 12
US-08-749-391-1
; Sequence 1, Application US/08749391
; Patent No. 5948667
; GENERAL INFORMATION:
; APPLICANT: Cheng, Kuo-Joan
```

APPLICANT: Selinger, Leonard B.
APPLICANT: Liu, Jin-Hao
APPLICANT: Hu, Youji
APPLICANT: Forsberg, Cecil W.
APPLICANT: Moloney, Maurice M.
TITLE OF INVENTION: A Xylanase Obtained From an
TITLE OF INVENTION: Anaerobic Fungus
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
City: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80803
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,391
FILING DATE: 13-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Donna M. Feider
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 93-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2058 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neocallimastix patriciarum
STRAIN: 47
IMMEDIATE SOURCE:
LIBRARY: genomic DNA library
CLONE: pmspx-06
FEATURE:
NAME/KEY: CDS
LOCATION: 301..1755
US-08-749-391-1
Query Match
Best Local Similarity 9.1%; Score 74.2; DB 2; Length 2058;
Matches 168; Conservative 0; Mismatches 138; Indels 1; Gaps 1;
DB 185 TGTGAAGATTATGTTATGGGATTACTAATTTCTAGATATATGTTGGTGAAGTT 244
DB 1747 TGTCAAAATTAATGATTGATTAATCAATTAATCAATTAATTTATTTGATTAATTAAT 1806
DB 245 AATATATATATTTCAATCCCTGTTTTATTTTTTTTATATAACAAATTAATAAT 304
DB 1807 AATTAAGAAAAAATTTTTTTTATTTTTTTTCTCTCAATTAATAA 1866
DB 305 TTAATTAATTTGGTTGTTTCAATTTATGACGGTAAATTTTATGTTAATTAATGAGC 364
DB 1867 TCATTAATAATAGATCATTAATTAATTAATTTTCAATTTTTTTTTTTTATTAATA 1926
DB 365 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 424
DB 1927 CGTAATAGTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1985
DB 425 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 484

DB 1986 CTAATTTAATAAATTATTAATAAAAAAAAAATATATAAAAAAAAAATATAAAAAAAAAA 2045
DB 485 TATTAAT 491
DB 2046 TATTAAT 2052
RESULT 13
US-09-390-200-1
Sequence 1, Application US/09390200
Patent No. 6137032
GENERAL INFORMATION:
APPLICANT: Cheng, Kuo-Juan
APPLICANT: Selinger, Leonard B.
APPLICANT: Liu, Jin-Hao
APPLICANT: Hu, Youji
APPLICANT: Forsberg, Cecil W.
APPLICANT: Moloney, Maurice M.
TITLE OF INVENTION: A Xylanase Obtained From an
TITLE OF INVENTION: Anaerobic Fungus
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
City: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80803
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/390,200
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/749,391
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Donna M. Feider
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 93-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2058 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Neocallimastix patriciarum
STRAIN: 27
IMMEDIATE SOURCE:
LIBRARY: genomic DNA library
CLONE: pmspx-06
FEATURE:
NAME/KEY: CDS
LOCATION: 301..1755
US-09-390-200-1
Query Match
Best Local Similarity 9.1%; Score 74.2; DB 3; Length 2058;
Matches 169; Conservative 0; Mismatches 138; Indels 1; Gaps 1;
DB 185 TGTGAAGATTATGTTATGGGATTACTAATTTCTAGATTAATGTTGGTGAAGTT 244

Gencore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2004, 20:34:55 ; Search time 3462 Seconds

(without alignments)
10216.036 Million cell updates/sec

Title: US-10-089-557-2_COPY_1_816

Perfect score: 1 actatagggcagcgtgtgc.....tatttcagttatgttaaaag 816

Sequence: 1 actatagggcagcgtgtgc.....tatttcagttatgttaaaag 816

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: 1: gb_ha: 2: gb_hg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vi: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_om: 21: em_or: 22: em_ov: 23: em_pat: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sts: 28: em_un: 29: em_vi: 30: em_hg_hum: 31: em_hg_inv: 32: em_hg_other: 33: em_hg_mus: 34: em_hg_pla: 35: em_hg_rod: 36: em_hg_mam: 37: em_hg_vrt: 38: em_sy: 39: em_hgo_hum: 40: em_hgo_mus: 41: em_hgo_other:

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	816	100.0	3041	6	AX370648
2	124.6	15.3	8056	6	AX599046 Sequence
3	123	15.1	8056	6	AX598900 Sequence
4	113	13.8	76568	3	MBREV
5	107.4	13.2	110000	2	PFMAL8pl_06
6	105.2	12.9	96799	9	AC104647
7	105.2	12.9	250029	3	AE014839
8	104.8	12.8	85779	8	SCE011856
9	104.4	12.8	47972	2	HSBA374F1
10	103.2	12.6	185596	6	AX599046
11	103.2	12.6	8056	6	AC021553
12	102.6	12.6	2009	6	AX457067
13	102.2	12.5	67970	3	PFMAL1P3
14	102.2	12.5	245802	2	AC006279
15	101.8	12.5	245802	2	AE014852
16	101.8	12.5	260929	3	AE014852
17	101.4	12.4	726	3	AB084758
18	100.8	12.4	76568	3	MBREV
19	100.8	12.4	114276	3	AC011355
20	100.8	12.4	166706	2	HSBA268B6
21	100.6	12.3	104992	2	AC005504
22	100.6	12.3	169546	2	AC004157
23	100.6	12.3	250431	3	AE014849
24	100.4	12.3	246611	2	AC111404
25	100.2	12.3	51552	2	PFMAL7P1_13
26	99.4	12.2	146570	3	AC117072
27	99.2	12.2	25568	3	AC116955
28	99	12.1	340552	3	PEA929354
29	98.8	12.1	86826	3	PFMAL3P5
30	98.6	12.1	1496	3	CEY53C12D
31	97.8	12.0	8056	6	AX598900
32	97.8	12.0	124635	9	AP000593
33	97.6	12.0	53932	2	AC023371
34	97.6	12.0	175544	2	AC117342
35	97.4	11.9	141701	2	AL671435
36	97.4	11.9	160401	9	AC009277
37	97.2	11.9	185596	9	AC021553
38	97.2	11.9	258658	3	AE014832
39	97	11.9	1496	3	CEY53C12D
40	97	11.9	8275	8	AY075133
41	97	11.9	8310	6	BD244527
42	97	11.9	203162	2	EX890546
43	96.8	11.9	258658	3	AE014832
44	96.8	11.9	144493	9	AP001547
45	96.8	11.9	181792	9	AC098822

ALIGNMENTS

RESULT 1
AX370648
LOCUS AX370648 3041 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 2 from Patent WO0210413.
ACCESSION AX370648
VERSION AX370648.1 GI:19168813
KEYWORDS
SOURCE
ORGANISM Arabidopsis sp.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Li,X., Cai,L., Cheng,N. and Liu,J.W.
TITLE Isolation and characterization of a fiber-specific actin promoter

Pred. No. is the number of results predicted by chance to have a

JOURNAL from cotton
Patent: WO 0210413-A 2 07-FEB-2002;
Institute of Molecular Agrobiolology (SG)
Location/Qualifiers
1. .3041
/organism="Arabidopsis sp."
/mol_type="unassigned DNA"
/db_xref="taxon:29726"

ORIGIN

Query Match 100.0%; Score 816; DB 6; Length 3041;
Best Local Similarity 100.0%; Pred. No. 1.4e-88;
Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ACTAATAGGCGCGCGTGTGCGAGCGCCGCGCTGCTCTTAAGAACAAATGTGTCAAG 60
DB 1 ACTAATAGGCGCGCGTGTGCGAGCGCCGCGCTGCTCTTAAGAACAAATGTGTCAAG 60
QY 61 TGGTCTTCCGCGAGCAATCCGAAATGAGGCTTAGAGTAACTTAACAGCGGACTGCT 120
DB 61 TGGTCTTCCGCGAGCAATCCGAAATGAGGCTTAGAGTAACTTAACAGCGGACTGCT 120
QY 121 CCAGCATTAAGTGTGCGAAATGTATGAGTGTGCTATGTCGAGATTTTCATG 180
DB 121 CCAGCATTAAGTGTGCGAAATGTATGAGTGTGCTATGTCGAGATTTTCATG 180
QY 181 AAGGTGTGAAGTAAATGATGATGAGTAACTTAATTTCTAGATTAATTTGGTTGGA 240
DB 181 AAGGTGTGAAGTAAATGATGATGAGTAACTTAATTTCTAGATTAATTTGGTTGGA 240
QY 241 AGTTAATATATATATATATATATATATATATATATATATATATATATATATAT 300
DB 241 AGTTAATATATATATATATATATATATATATATATATATATATATATATATAT 300
QY 301 TAAATTTAATTAATTTGTTGTTTCAATTTTGAAGGTTAAATTTTGAATTAAT 360
DB 301 TAAATTTAATTAATTTGTTGTTTCAATTTTGAAGGTTAAATTTTGAATTAAT 360
QY 361 GAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
DB 361 GAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
QY 421 AATAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
DB 421 AATAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 481 AATATATATATATATATATATATATATATATATATATATATATATATATAT 540
DB 481 AATATATATATATATATATATATATATATATATATATATATATATATATAT 540
QY 541 TTGCAAGTTCACCAACACGAGAAAGTAATCAATTTTCAGAAAAGTAATCAATTTTC 600
DB 541 TTGCAAGTTCACCAACACGAGAAAGTAATCAATTTTCAGAAAAGTAATCAATTTTC 600
QY 601 AGAATTAATTTTTCGAAATTAATTTTACGCAACAAATGAGTCTAAGTCTCTGTT 660
DB 601 AGAATTAATTTTTCGAAATTAATTTTACGCAACAAATGAGTCTAAGTCTCTGTT 660
QY 661 TTTATTTTATTTTTCATTTTAGAGAAATAGAAATGATTTGCAAAATGCTTTAATC 720
DB 661 TTTATTTTATTTTTCATTTTAGAGAAATAGAAATGATTTGCAAAATGCTTTAATC 720
QY 721 TAGCTTTGTTAGATTAAGTGAAGGCGACAGAACCCGCGTTGCAAGAGATTTTGCCTGAC 780
DB 721 TAGCTTTGTTAGATTAAGTGAAGGCGACAGAACCCGCGTTGCAAGAGATTTTGCCTGAC 780
QY 781 TCACCTACCTAGATTTCTATTTTCAGATTTGTAAGG 816
DB 781 TCACCTACCTAGATTTCTATTTTCAGATTTGTAAGG 816

```

RESULT 2
AX599046

LOCUS AX599046 8056 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 386 from Patent WO02077272.
ACCESSION AX599046
VERSION AX599046.1 GI:28399186
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.

REFERENCE
1
AUTHORS Berlin, K., Braun, A., Disler, J., Guerig, D., Howe, A., Mueller, J.,
Oler, A., Piepenbrock, C., Adorjan, P., Grabs, G., Lesche, R., Ley, E.,
Lewin, A., Lipscher, E., Maier, S., Model, F., Mueller, V., Otto, T.,
Pele, C., and Ziebarth, H.
TITLE Methods and nucleic acids for the analysis of hematopoietic cell
proliferative disorders
JOURNAL Patent: WO 02077272-A 386 03-OCT-2002;
EpiGenomics AG (DE)

FEATURES

source

Location/Qualifiers
1. .8056
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

ORIGIN

Query Match 15.3%; Score 124.6; DB 6; Length 8056;
Best Local Similarity 53.3%; Pred. No. 2.2e-06;
Matches 285; Conservative 0; Mismatches 249; Indels 1; Gaps 1;

```

QY 189 AAGATTAATGATTAATTTGGGATTAATTTCTAGATTAATTTGAGTTGGAAGTAATA 248
DB 1646 AAAAAATTTTAAATTAATTAATTAATTAATTAATTTTAAATTAATTAATTAATTAAT 1705
QY 249 TATAATTAATCAATCTGTTTAAATTTTAAATTTTAAATTAATTAATTAATTAATTAAT 308
DB 1706 TATAATTAATTAATTAATTAATTTTAAATTTTAAATTTTAAATTAATTAATTAATTT 1765
QY 309 TTAATTTGTTGTTTCAATTTTAAATTAATGACGGT-TAAATTTTAAATTAATTAATTAAT 367
DB 1766 TTAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1825
QY 368 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 427
DB 1826 AATTTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 1885
QY 428 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 487
DB 1886 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1945
QY 488 AATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTGCAGG 547
DB 1946 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2005
QY 548 TTCAATCCAAACCCGAAAGTAATCAATTTTCAGAAAAGTAATCAATTTTCAGAAAT 607
DB 2006 AATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAAAT 2065
QY 608 AATTTTCGAAATTAATTTTACGCAACAAATGAGTCTAAGTCTCTGTTTAAAT 667
DB 2066 TTATTTAAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTTAA 2125
QY 668 TTTATTTTCTATTTTAGAGAAATGAGAAATGATTTGCAAAATGCTTTAATCTA 722
DB 2126 TAAATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTTTA 2180

```

RESULT 3
AX598900 8056 bp DNA linear PAT 14-FEB-2003
LOCUS AX598900
DEFINITION Sequence 240 from Patent WO02077272.
ACCESSION AX598900
VERSION AX598900.1 GI:28399038
KEYWORDS

SOURCE	ORGANISM	synthetic construct
REFERENCE	AUTHORS	artificial sequences.
TITLE	Berlin, K., Braun, A., Distler, U., Guebig, D., Howe, A., Mueller, U., Olk, A., Piepenbrock, C., Adorjan, P., Grabs, G., Iesche, R., Leu, B., Lewin, A., Lipscher, E., Maier, S., Model, F., Mueller, V., Otto, T., Pelet, C. and Ziebarth, H.	Method and nucleic acids for the analysis of hematopoietic cell proliferative disorders
JOURNAL	Patent: WO 0207272-A 240 03-OCT-2002;	Epigenomics AG (DE)
FEATURES	source	Location/Qualifiers
		1..8056
		/organism="synthetic construct"
		/mol_type="unassigned DNA"
		/db_xref="taxon:32630"
		/note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN		
Query Match	15.1%; Score 123; DB 6; Length 8056;	
Best Local Similarity	53.1%; Pred. No. 3.5e-06;	
Matches 284; Conservative	0; Mismatches 250; Indels 1; Gaps 1;	
QY	189 AAGAATTACGTTTGGGATTCTCAATTCTAGATTATTTGGTTGGAGTTATA	248
Db	1646 AAAAAATATTTAATTAATACGATTAATTTTATTTAATTAATTAATTAATTT	1705
QY	249 TATAATATTCATCCCTGTTTATTTTATTTTATTAACACATTCATATATTAT	308
Db	1706 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1765
QY	309 TTAACCTTGGTTGTTTCATTTATGACGGT-TAATATTTAGTTAATTAATGACAT	367
Db	1766 TTAATTAATAAATAATTAATAATTAATTAATTAATTAATTAATTAATAAATAA	1825
QY	368 ATTATATTTTAAATTAATTAATCATTTGTAATATGTAATAATTAATTAATTAAT	427
Db	1826 ATTTGGTTTAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAAT	1885
QY	428 TATTAATTAATTAATTAATTAATCAATCAACAATTAATAATTAATTAATTAAT	487
Db	1886 TTTTATTAATTAATAAATTAATAATTTTAATAAATAAATAAATAAATAAATAAAT	1945
QY	488 AATATTTTAAATAGCTTTTCAGTAATCTGTCAACATAGAAATATTTTTCAGG	547
Db	1946 AAAAAAATATTAATTTTAAATAAATAATTAATTAATTAATTTTATTAATAATTA	2005
QY	548 TTCAATCCAAACCAAGAAAGTAATCACTTTTCAGAAAGTAATCAATTTTCAGAAAT	607
Db	2006 AATATTTTAAATTAATTTTAAATAAATAAATAAATAAATAAATAAATAAAT	2065
QY	608 ATTTTGGGAATTAATTTTACGCGAACAATGAGCTCAAGTTCCTGTTTAAAT	667
Db	2066 TTAATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2125
QY	668 TTTATTTTTCATTTAGAGAACTGGAATGATTTGCAATGCTTTAACTGA	722
Db	2126 TAATTTTATTTATTAATTAATAAATAAATAAATAAATAAATAAATAAATTAAT	2180
RESULT 4		
MBREV/c	76568 bp	DNA
LOCUS	Monosiga brevicollis	mitochondrion, complete genome.
DEFINITION	AF538053 AF275274	
ACCESSION	AF538053.1	GI:25344058
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		

	TITLE
JOURNAL	A novel motif for identifying rps3 homologs in fungal mitochondria
MEDLINE	trends Biochem. Sci. 25 (8), 363-365 (2000)
PUBMED	20377911
REFERENCE	2 (bases 1 to 76568)
AUTHORS	Lang,B.F., O'Kelly,C., Neraud,T., Gray,M.W. and Burger,G.
TITLE	The closest unicellular relatives of animals
JOURNAL	Curr. Biol. 12 (20), 1773-1778 (2002)
MEDLINE	22288938
PUBMED	12401173
REFERENCE	3 (bases 64800 to 65249)
AUTHORS	Lang,F.B. and Bullerwell,C.
JOURNAL	Direct Submission
TITLE	Submitted (06-JUN-2000) Biochimie, Université de Montreal, 2900
REFERENCE	Boul. Edouard Montpetit, Montreal, QUE H3C 3J7, Canada
AUTHORS	4 (bases 1 to 76568)
TITLE	Burger,G. and Lang,B.F.
JOURNAL	Direct Submission
TITLE	Submitted (10-AUG-2002) Biochimie, Université de Montreal, 2900,
COMMENT	Boul. Edouard-Montpetit, Montreal, Quebec H3T 1J4, Canada
FEATURES	On Sep 29, 2002 this sequence version replaced gi:9966505.
SOURCE	Location/Qualifiers
	1..76568
	/organism="Monosiga brevicollis"
	/organella="mitochondrion"
	/mol_type="genomic DNA"
	/strain="ATCC 50154"
	/db_xref="ATCC:50154"
	/db_xref="taxon:61824"
gene	17..2894
	/gene="rnl"
rRNA	17..2894
	/gene="rnl"
gene	/product="large subunit ribosomal RNA"
	4627..6123
gene	/gene="nad4"
	4627..6123
CDS	/EC number="1.6.5.3"
	/codon_start=1
	/transl_table=4
	/product="NADH dehydrogenase subunit 4L"
	/protein_id="AAZ8335.1"
	/db_xref="GI:23344059"
	/translation="MMNTLLPFPALPFIGILLILPEFNNELNIKKGQIALGTSIILL
	LISLRIMFHEDYTSKROYILKLNMINDELTSNMIVLGIDSLVFILTLLEF
	ICLSIRMDSVKIILVKGYSIFRVEILLIDVFETDIFGYILLFEGLIPMYLIIGIM
	GSRQKTAGTYFFPYFTLTSSVMLEILVHSISTGTDTVLITGYRIDPRACKYTEL
	AFPAFLAVLKPKPEPHMLPQAHEVAENVASVILIAGLIKGGYGFTRFTPLLPDMS
	HYPAFLVILGVLAIVASLTITLRQDLKIKIAVSSSHGLVSLIFPTLANIMSGS
	IPLQALGVASSALFIIVLVLYDRHLRIKVYSGMTGWPIVSFLFPLANIAYVAF
	LSGNCEVEFLCLAIPEVNVTVAIIIASGLIISACSYLFLYNRVCFSFYLSNMSY
	HLSIKRSRRREFVLEPVYLITLITGVYHPHFIDLNISVINLTPL"
gene	7006..7089
	/gene="trnL"
cRNA	7006..7089
	/gene="trnL"
	/product="tRNA-Leu"
gene	7386..7736
	/gene="rps14"
CDS	7386..7736
	/gene="rps14"
	/codon_start=1
	/transl_table=4
	/product="ribosomal protein S14"
	/protein_id="AAN28336.1"
	/db_xref="GI:23344060"
	/translation="MNRIKLKFETHSKLDPFKKRTLYVKNENNHITYNALIKQIIDL
	SVKNSIAEGSISEYFDNKDPLSLIPSHIRNCKLTGRSRGNLPKYGLSRIMPETELA
	SEGIIIPGMATYKNKT"
	8034..8513
gene	/gene="rps8"

CDS
8034..8513
/gene="rps8"
/codon_start=1
/transl_table=4
/product="ribosomal protein S8"
/protein_id="AAN28337.1"
/db_xref="gi:23344061"
/translation="MNELISKNNNTYRKMLTCQPNNDYSLIKVLSKSEGVILLET
NKNKQKILNFKLNDKLGKLVKPKIPKPGKXYISNKEILMSNTKKGFTINT
SKIGKILYSNLSKHKYILRTSSNNILQNNKGIITQEARIKNNVGGELLVYKFC"
10529..11287
/gene="atp6"
10529..11287
/gene="atp6"
/EC_number="3.6.1.34"
/codon_start=1
/transl_table=4
/product="ATP synthase F0 subunit 6"
/protein_id="AAN28338.1"
/db_xref="gi:23344062"
/translation="MAYGAPDQFELIRIIPHLFGNLDISTNSTIEMFIALIYLIV
YVSGIKESGLVPSWQSMELIYEYFGVEDNKQKESQAPPMILYEMFIVM
NLFGIPEYFSPTAHIAVTFGLSISIFGLVGNHNSYFSMFMAGSLIIAP
MIVELISHTKAVSIGRLAANTAGHILFALISGFTWMLTSGFMIIASLPMAL
VLFITIMAVAVIOAVVFSLLTALYSESHLH"
11668..11967
/gene="atp8"
11668..11967
/gene="atp8"
/EC_number="3.6.1.34"
/codon_start=1
/transl_table=4
/product="ATP synthase F0 subunit 8"
/protein_id="AAN28339.1"
/db_xref="gi:23344063"
/translation="MPQLSTEFVSCQWPELVILVEYLTISNVPRIAIKIKTRK
CMDESIYTISSDIQTLNNENPAQIINVINSVSNHNTIVSSPENANKWIKK"
13729..14760
/gene="nad1"
13729..14760
/gene="nad1"
/EC_number="1.6.5.3"
/codon_start=1
/transl_table=4
/product="NADH dehydrogenase subunit 11"
/protein_id="AAN28340.1"
/db_xref="gi:23344064"
/translation="MTNLILITLITLIIKILIAIVPVULIAYVLIARXYIASMOGR
KGPNAVGVFGILOPLADSKLFEKNAVLPHTANMIIFVSVVAAPFTLALIMAVIPYN
EGKVIADVNIQLYIPAVSSISVAIIMSQASNSKAFPSAIPRAAAMISYEWSGL
ILSVILCVGGINITISIVLAQKNIWFILPEPAPIMFVSALAEINRAPDLTSESE
LVSGNAYVSAIFTLFELIAYTHIIFMSITLIFEGMWLPPPEPFIFNLPSQIML
ALTKGFIIFVFWVWRAFPFRPRVYMLMELMKSYLPSLGLIITASIIAMFNVOI
SKFINLIIY"
15143..15214
/gene="trn1"
15143..15214
/gene="trn1"
/product="tRNA-Thr"
15536..15617
/gene="trn5"
15536..15617
/gene="trn5"
15536..15617
/gene="trn5"
15742..15816
/product="tRNA-Ser"
15742..15816
/gene="trn5"
15742..15816
/product="tRNA-His"
16105..16500
/gene="rps12"
16105..16500
/gene="rps12"
CDS

/codon_start=1
/transl_table=4
/product="ribosomal protein S12"
/protein_id="AAN28341.1"
/db_xref="gi:23344065"
/translation="MTTINQSKNNIRKRRITRITAGSKFLNKGILNPQLKGVCIKRL
TRKPKPNSAQKRIAKVLTNGLLVDVAFIPSEHSLDGHGIVLIRGGVKKDLPGLK
CVRKQYDLAAGVIGKTSRSKYGKPKYK"
17571..18104
/gene="rpl16"
17571..18104
/gene="rpl16"
/codon_start=1
/transl_table=4
/product="ribosomal protein L16"
/protein_id="AAN28342.1"
/db_xref="gi:23344066"
/translation="MLNKRPERKTKYKQKRIKINSNLPYKNSNLFQKULENKDPRYLL
IAMEPLRITASHIAGELAIKRLNLIKPKSSNLSITDIEKRNINKQYNIHVPFLP
VTKRPABRMKGKGSIDVWTRIRIGSPFELKFKANINNKNIIQIRCNPIAKQIF
DVNSKIGIKTKIKYKVS"
19351..19422
/gene="trn8"
19351..19422
/gene="trn8"
/product="tRNA-Arg"
19569..23450
/gene="nad5"
19569..23450
/gene="nad5"
/EC_number="1.6.5.3"
/codon_start=1
/transl_table=4
/product="NADH dehydrogenase subunit 5"
/protein_id="AAN28343.1"
/db_xref="gi:23344067"
/translation="WYLIILPLPLSLVIAAGLGRFPGTKAGIITTCIMTSLSEW
IAFEEVGINSPVYIPLMKMINVGTIYDMSGLMYDSITVMTLILVTVSSLVHYSI
SYMDGDPLHPRFMSYLSLFTFSVMYLTADSYLLCVGWSVGCYLLINFWYRIQ
ANKSAIKAMWNRVGDMLFLGMPAIFVEFTLIDYSTRFSYKLSDEYTIIFENYRIH
ALTLISILFLGAGKSAOLGHTMLDAMGPPPVSAIIHAAMVTGAVGLLIRSSP
IIEYAPKAVIITITVGTITLPAISIGIIONDIKKVIASTGQSGVAVLACISNVA
VALFHLTHGFRKALLFLGAGSVTHAMGDEDPKRLGGLIRIRIPTYSMMLIGSLIM
GFPPLAGYSKALLLEIAYGKYITIDGPAHLGLIAGMVAIVSIRLIMYFLSDTNG
Query Match 13.8%; Score 113; DB 3; Length 76568;
Best Local Similarity 51.0%; Pred. No. 3.2e-05;
Matches 294; Conservative 0; Mismatches 280; Indels 3; Gaps 1;
QY 169 GPAATTCAGGAAAGTGTGAAGATTGATGTTATGCGATTACTAATTTCTAGATTAA 228
DB 38187 GTGATATTAAGGAAATTTTATTATTAATCATTTGTTAATAAATCAATTCGATTTTAA 38128
QY 229 TTGTGTTTGAAGTATATATAATTAATTCATCCCTGTTTATTTTATTTTATTA 288
DB 38127 CCAATCTTAAWAAACATTTTATTATTAATTTGTTTTTATTTTATTAATAAAAAA 38068
QY 289 CACAAATTAACAATATTATTTAACTTTGTTTCAATTTAAGCGTTAATATT 348
DB 38067 TATATTTATTTATTTTCAATATTAATTTAATTTATTTATTAATTAATAAAAAA 38008
QY 349 AGTTAATATGAGCATTTATTAATTAATAAATAAATAATCAATGTAATAATGATAAA 408
DB 38007 AAAATACAAATATTAATTAATTAATAAATAATTAATTAATTAATAAATAATTAATAA 37948
QY 409 AATAATTAATTAATTAATTTAATAATAATAATAATAAATCAATCAATCAATTAATAAAGAT 468
DB 37947 TAAATATTAATTAATTTAATAATAATAATAATAATAATAATAATAATAATAAT 37868
QY 469 AATAAATCTTAATTAATTAATTTTAAATAGCTTTTCAGTAATCTGTCAACAAAT 528
DB 37867 AATATTAATTTAAATAATAATTAATTTTAAATAATAATAATAATAATAATAATTTAA 37828

```

Ox      529  AGAAATATTGTTTCAGGTGCATCCAAACACGAAAAAGTAATCATTTTTCAGAAAAGT   588
Db      37827 AATAATTTATTTTAAAAAATATATATATATATATATATATATATATATATATATAT   37668
Ox      589  AA--ATCATTTTTCAGAAATATTTTTTCGAAATATATTTTACTGGCAACAATGGAGT   645
Db      37767 AAAATATATATATATATATATATATATATATATATATATATATATATATAT   37708
Ox      646  CTAAGTGTTCTCGTTTATTTTATTTTATTTTCTATTTTAGAGAACTAGAAATTTGATT   705
Db      37707 TTATGTACAACTTTTCTTTTTCGAAATCTTTTACAAATGAGACCCTTAACTGTAC   37648
Ox      706  CAATGTCTTATATCTAGCTGTGAATAGATTTAGATGAA   742
Db      37647 AATAGAGATATATCATTTATTTTATATATATATA   37611

RESULT 5
PFMALBp1_06
WPCOMMENT
Sequence split into 14 fragments LOCUS PFMALBp1 Accession AL844507
Fragment Name Begin End
PFMALBp1_00 1 11000
PFMALBp1_01 10001 21000
PFMALBp1_02 20001 31000
PFMALBp1_03 30001 41000
PFMALBp1_04 40001 51000
PFMALBp1_05 50001 61000
PFMALBp1_06 60001 71000
PFMALBp1_07 70001 81000
PFMALBp1_08 80001 91000
PFMALBp1_09 90001 101000
PFMALBp1_10 100001 111000
PFMALBp1_11 110001 121000
PFMALBp1_12 120001 131000
PFMALBp1_13 130001 1325595
Continuation (7 of 14) of PFMALBp1 from base 600001 (AL844507 Plasmodium falciparum 3D)

Query Match 13.2%; Score 107.4; DB 2; Length 110000;
Best Local Similarity 49.2%; Pred. No. 0.00013;
Matches 282; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

Ox      167  AACGATTTTCATGGAAGGTGTAGCAATTAATGTTTATGCGATTACTAATTTCTAGATT   226
Db      44209 AATATTTTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTCTA   44268
Ox      227  AATGCGTTTGGAGTTATATATATATATTCATCCGTGTTTATTTTATTTTTTTAT   286
Db      44269 AATAGTATTTTTTTTAAAAAACCAATTTCAACAAGAATTTTAAAAANTTTATTTATATAA   44328
Ox      287  AACACAAATTCACAAATATTTATTTTAATCTTGTTTCAATTTNAGACGTTAATATT   346
Db      44329 AATATATATATATATATATATATATATATATATATATATTTATTTATTTATATATATA   44388
Ox      347  TTAGTTTAATAATTGAGCATTTATATATATTAATAATAATCAATGTAATATATGTAA   406
Db      44389 TAATTAATTAATTAATTAATTAATTTTATTTATTTAGCAAAAGAAATTTTTTAATTTAT   44448
Ox      407  AATAATTTAAAAATATATATTTATATATATATATATATATATATATATATATATAT   466
Db      44449 AAAAAATATATATATATATATATTTTATTTTAAAAACAAAAAATTTTTTTTTTTTTTTT   44508
Ox      467  ATATATTAATTTCTTAATATATATATTTTTTAAATATAGCTTTCAGTAATCTGTCAA   526
Db      44509 AATTTTTTTTATATATATATATATTTTCATATATATATATATATATATATATATAT   44568
Ox      527  ATGGAATAATTTTTTTCAGGTTCATCCAAACACGAAAAAGTAATCATTTTCAGAAA   586
Db      44569 AAAAAAAGAAATGTTCTATATAAAAAACCCGTTGTTTTTATTTTGAATATTTAAAAATAT   44628
Ox      587  GTAAATCATTTTCAGAAATATTTTTTCGAAATATTTTATACGTGCGAAACAAATGAG   646
Db      44629 ATGCATTAATAATTTTTTTTTTTTTTTTTTTTTTTTAAAGATTCAAAAATATAATTTTAG   44688

```

```

OY      647  TAAGTGGTTCGTTTATTTATTTATTTTTCATTTTATGAGAAACCTGAAATTCATTGC 706
Db      44689  TATTTTATAGTACTAACATATTTTATTTATTTTATATATTTATTTTATTTTGAATTTTA 44744
OY      707  AAATGCTTTATATCTAGCTTGTTTACATTAGTT 739
Db      44749  AAATATTTATATATTTTATTTTATTTTATTTT 44781

RESULT 6
AC104647      96799 bp      DNA      linear      PRI 21-FEB-2002
LOCUS      AC104647
DEFINITION      Homo sapiens BAC clone RP11-38D12 from 4, complete sequence.
ACCESSION      AC104647
VERSION      AC104647.3      GI:18464257
KEYWORDS
SOURCE      HTG.
ORGANISM      Homo sapiens (human)
Homosapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Sulston,J.E. and Waterston,R.
TITLE      1 (bases 1 to 96799)
JOURNAL      Toward a complete human genome sequence
MEDLINE      Genome Res. 8 (11), 1097-1108 (1998)
PUBMED      99063792
PUBMED      9847074
REFERENCE
AUTHORS      2 (bases 1 to 96799)
TITLE      Sweatengen-Shahid,S., Haakenson,W. and Creason,K.
JOURNAL      The sequence of Homo sapiens BAC clone RP11-38D12
REFERENCE
AUTHORS      3 (bases 1 to 96799)
TITLE      Waterston,R.H.
JOURNAL      Direct Submission
Submitted (18-DEC-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 96799)
Waterston,R.H.
Direct Submission
Submitted (01-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 96799)
Waterston,R.
Direct Submission
Submitted (21-FEB-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb 1, 2002 this sequence version replaced gi:18071455.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wuston.wustl.edu
----- Summary Statistics
Center project name: H_MK0038D12
Drafting Center: WIBR
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

```

MAPPING INFORMATION:

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.edu
 ----- Summary Statistics -----
 Center project name: H.NH003812
 Drafting Center: WIBR

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS	1 (bases 1 to 96799)
JOURNAL	Sulston,J.E. and Waterston,R.
MEDLINE	toward a complete human genome sequence
PUBMED	Genome Res. 8 (11), 1097-1108 (1998)
REFERENCE	99063792
AUTHORS	9847074
TITLE	2 (bases 1 to 96799)
JOURNAL	Sweatengen-shanid,S., Haakenson,W. and Creason,K.
REFERENCE	The sequence of Homo sapiens BAC clone RP11-38D12
AUTHORS	Unpublished (2001)
TITLE	3 (bases 1 to 96799)
JOURNAL	Waterston,R.H.
REFERENCE	Direct Submission
AUTHORS	Submitted (18-DEC-2001) Genome Sequencing Center, Washington
TITLE	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL	MO 63108, USA
REFERENCE	4 (bases 1 to 96799)
AUTHORS	Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (01-FEB-2002) Genome Sequencing Center, Washington
REFERENCE	University School of Medicine, 4444 Forest Park Parkway, St. Louis,
AUTHORS	MO 63108, USA
TITLE	5 (bases 1 to 96799)
JOURNAL	Waterston,R.
REFERENCE	Direct Submission
AUTHORS	Submitted (21-FEB-2002) Department of Genetics, Washington
TITLE	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
JOURNAL	On Feb 1, 2002 this sequence version replaced gi:18071455.
COMMENT	

RESULT	6
LOCUS	AC104647
DEFINITION	AC104647 Homo sapiens BAC clone RP11-38D12 from 4, complete sequence.
ACCESSION	AC104647
VERSION	AC104647.3 GI:18464257
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

QY	Db	QY	Db
647	TAAGTGTCTGTGTTTATTTTATTTTTCATTTAGAGAAACTGAAATTTGATTTGC	707	AAATGCTTTAATCTAGCTGTGTAAGT
44689	TATTTAATAGTACATTTTTTATTTATTTTAAATATTTATTTTATTTGATTTTA	44743	AAATATTTATATAATTTTTTTTTTTTTTTT
			44781

1

Page 1816.rqe

1

REFERENCE	1	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
AUTHORS	Foury, F.; Roganti, T.; Lecrenier, N. and Purnelle, B.	
TITLE	The complete sequence of the mitochondrial genome of Saccharomyces cerevisiae	
JOURNAL	FEBS Lett. 440 (3), 325-331 (1998)	
MEDLINE	99087401	
PUBMED	9872396	
REFERENCE	2 (bases 1 to 85779)	
AUTHORS	MIPS.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-DEC-1998) Data collected by MIPS on behalf of the European yeast mitochondrial genome sequencing project. MIPS at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152 Martinsried, Germany; E-mail: Mewes@mips.biochem.mpg.de	
FEATURES	Location/Qualifiers	
source	1..85779 /organism="Saccharomyces cerevisiae" /organalle="mitochondrion" /mol_type="genomic DNA" /strain="FY1679, isogenic derivative of S288C" /db_xref="taxon:4932" 1..11 /product="9S ribosomal RNA" 731..802 /gene="rP(TGG)Q - systematic name" 731..802 /gene="rP(TGG)Q - systematic name" /product="rRNA-Pro - Common name" /note="rRNA-Pro - Common name anticodon gene: TGG" 3952..4338 /note="unnamed protein product; ORF Q0010" /codon_start=1 /transl_table=3 /protein_id="CAA09822.1" /db_xref="GI:4160363" /db_xref="GOA:Q9ZZX9" /db_xref="SPTREMBL:Q9ZZX9" /translation="MYVIMPLVNMILITILIFYSIVGVPIIFNNNYDDPDIPLFII VYFKIITLIVLYVMTNYIVTPSSGSPRGRGVIILYNMLYSYNNFIDVYMKFIC VTYVIMLWMLSPITPSPIYSEVPS" complement(4012..4312) /note="ori1" 4254..4415 /note="unnamed protein product; ORF Q0017" /codon_start=1 /transl_table=3 /protein_id="CAA09822.1" /db_xref="GI:4160364" /db_xref="GOA:Q9ZZX8" /db_xref="SPTREMBL:Q9ZZX8" /translation="MCATYMPNITYIITHPTILRTGRGFVRNEDLIYKKXKNLIN NLNNMTYL"	
CDS	4254..4415 /product="15S ribosomal RNA" 9374..9447 /gene="rW(TCA)Q - systematic name" 9374..9447 /gene="rW(TCA)Q - systematic name" /product="rRNA-Tip1 /note="rRNA-Tip1 - common name anticodon gene: TCA" 11667..11957 /note="unnamed protein product; ORF Q0032" /codon_start=1 /transl_table=3 /protein_id="CAA09823.1" /db_xref="GI:4160365" /db_xref="GOA:Q9ZZX7" /db_xref="SGD:S0007259" /db_xref="SPTREMBL:Q9ZZX7"	
gene	rep_origin	
CDS	complement(12510..12780) /note="ori8" join(13818..13986,16435..16470,18954..18991,20508..20984, 21995..22246,23612..23746,25318..25342,26229..26701) /gene="COX1" join(13818..13986,16435..16470,18954..18991,20508..20984, 21995..22246,23612..23746,25318..25342,26229..26701) /gene="COX1" /note="ORF Q0045" /codon_start=1 /transl_table=3 /protein_id="CAA09824.1" /db_xref="GI:4160366" /db_xref="GOA:P00401" /db_xref="GOA:Q9ZZX6" /db_xref="SWISS-PROT:P00401" /translation="MVORMLYSTNAKDIAVLVFMALIFSGMAGTAMSLIRLELAAG SOYLHNSQSLPNLYVGHAVLMTFLVMTLIGSGFVYLPIMGADTAPRINNTA FVLPWGLVCLTSTLVESGAGTGWTPPLLSIOHSGSVDAIPALHITSLSLL GAINFIYTLTNKRTNEMTHKPLFVWSIFITAFLLLSLPVLSAGITMLLDNENT SFEVSGSGDPIIYEHLPFMRFGHEVYVLIIPGGIISHVATYSKRPVGEISMYVA MASIGLGFVWSHMYTVGLDADTRAFYSATWIIAIPGIIKIFSMLATIHGSIIL ATPMLVAIAFLPFLTMGLTGVALANSLVADPHDYVYVGHFVYLSMGAFISLPAG VYVWSPOILGLANVNEKIAOIFQPLIFGANVIFPMHRLINGMWRIPVDPDAFACW NYVASIGSFATISLFLPIYILYDQVNGLNKNKNSYVYNKAPDVESENITNNT VKSSISEFLITSPVASFNTPAVOS" join(13818..13986,16435..16470,18954..18991,20508..20984, 21995..23167) /gene="SCA15alpha" join(13818..13986,16435..16470,18954..18991,20508..20984, 21995..23167) /gene="SCA15alpha" /note="ORF Q0070" /codon_start=1 /transl_table=3 /protein_id="CAA09829.1" /db_xref="GI:4160371" /db_xref="SPTREMBL:Q9ZZX1" /db_xref="SPTREMBL:Q9ZZX1" /translation="MVORMLYSTNAKDIAVLVFMALIFSGMAGTAMSLIRLELAAG SOYLHNSQSLPNLYVGHAVLMTFLVMTLIGSGFVYLPIMGADTAPRINNTA FVLPWGLVCLTSTLVESGAGTGWTPPLLSIOHSGSVDAIPALHITSLSLL GAINFIYTLTNKRTNEMTHKPLFVWSIFITAFLLLSLPVLSAGITMLLDNENT SFEVSGSGDPIIYEHLPFMRFGHEVYVLIIPGGIISHVATYSKRPVGEISMYVA MASIGLGFVWSHMYTVGLDADTRAFYSATWIIAIPGIIKIFSMLNPFSDXKN NKNKKLIRNYOKNNNNMMKTYLNNNNMIMMNVKGLYDIPRSNNRYIOPNINKE LVYGVNLESVCMPYTYNIVKHWVGIENNLIYMTGIIITLDGIDVTSKDDLKKTI MEINCRFLRKOSMTHSEYLMIVFVMLLSHYCHMSYPRKIAKYKGSYNQLEPYTSLLPC FTILIRYFNGRVKIVPNNLYDLNLYSLAHMTCDSPVNGGLVYNLOSFTTKELI FINMLIKIKFNLCITLHRSRNNKYTIYKRVESVKRLPFIYKYIILPSMKYKEDIMLMOK KYNMIN"	
gene	gene	
CDS	join(13818..13986,18954..18991,20508..21935) join(13818..13986,18954..18991,20508..21935) /gene="I-SCB1" /note="I-SCB1" /note="ORF Q0065" /codon_start=1 /transl_table=3 /protein_id="CAA09828.1" /db_xref="GI:4160370" /db_xref="GOA:Q9ZZX2" /db_xref="SPTREMBL:Q9ZZX2" /translation="MVORMLYSTNAKDIAVLVFMALIFSGMAGTAMSLIRLELAAG SOYLHNSQSLPNLYVGHAVLMTFLVMTLIGSGFVYLPIMGADTAPRINNTA FVLPWGLVCLTSTLVESGAGTGWTPPLLSIOHSGSVDAIPALHITSLSLLGAINFIYTLTNKRTNEMTHKPLFVWSIFITAFLLLSLPVLSAGITMLLDNENT SFEVSGSGDPIIYEHLPFMRFGHEVYVLIIPGGIISHVATYSKRPVGEISMYVA MASIGLGFVWSHMYTVGLDADTRAFYSATWIIAIPGIIKIFSMLATIHGSIIL ATPMLVAIAFLPFLTMGLTGVALANSLVADPHDYVYVGHFVYLSMGAFISLPAG VYVWSPOILGLANVNEKIAOIFQPLIFGANVIFPMHRLINGMWRIPVDPDAFACW NYVASIGSFATISLFLPIYILYDQVNGLNKNKNSYVYNKAPDVESENITNNT VKSSISEFLITSPVASFNTPAVOS"	
CDS	join(13818..13986,16435..16470,18954..18991,20508..20984, 21995..23167) /gene="SCA15alpha" join(13818..13986,16435..16470,18954..18991,20508..20984, 21995..23167) /gene="SCA15alpha" /note="ORF Q0070" /codon_start=1 /transl_table=3 /protein_id="CAA09829.1" /db_xref="GI:4160371" /db_xref="SPTREMBL:Q9ZZX1" /db_xref="SPTREMBL:Q9ZZX1" /translation="MVORMLYSTNAKDIAVLVFMALIFSGMAGTAMSLIRLELAAG SOYLHNSQSLPNLYVGHAVLMTFLVMTLIGSGFVYLPIMGADTAPRINNTA FVLPWGLVCLTSTLVESGAGTGWTPPLLSIOHSGSVDAIPALHITSLSLL GAINFIYTLTNKRTNEMTHKPLFVWSIFITAFLLLSLPVLSAGITMLLDNENT SFEVSGSGDPIIYEHLPFMRFGHEVYVLIIPGGIISHVATYSKRPVGEISMYVA MASIGLGFVWSHMYTVGLDADTRAFYSATWIIAIPGIIKIFSMLNPFSDXKN NKNKKLIRNYOKNNNNMMKTYLNNNNMIMMNVKGLYDIPRSNNRYIOPNINKE LVYGVNLESVCMPYTYNIVKHWVGIENNLIYMTGIIITLDGIDVTSKDDLKKTI MEINCRFLRKOSMTHSEYLMIVFVMLLSHYCHMSYPRKIAKYKGSYNQLEPYTSLLPC FTILIRYFNGRVKIVPNNLYDLNLYSLAHMTCDSPVNGGLVYNLOSFTTKELI FINMLIKIKFNLCITLHRSRNNKYTIYKRVESVKRLPFIYKYIILPSMKYKEDIMLMOK KYNMIN"	
gene	gene	
CDS	join(13818..13986,16435..16470,18954..18991,20508..20984, 21995..23167) /gene="SCA15alpha" join(13818..13986,16435..16470,18954..18991,20508..20984, 21995..23167) /gene="SCA15alpha" /note="ORF Q0070" /codon_start=1 /transl_table=3 /protein_id="CAA09829.1" /db_xref="GI:4160371" /db_xref="SPTREMBL:Q9ZZX1" /db_xref="SPTREMBL:Q9ZZX1" /translation="MVORMLYSTNAKDIAVLVFMALIFSGMAGTAMSLIRLELAAG SOYLHNSQSLPNLYVGHAVLMTFLVMTLIGSGFVYLPIMGADTAPRINNTA FVLPWGLVCLTSTLVESGAGTGWTPPLLSIOHSGSVDAIPALHITSLSLL GAINFIYTLTNKRTNEMTHKPLFVWSIFITAFLLLSLPVLSAGITMLLDNENT SFEVSGSGDPIIYEHLPFMRFGHEVYVLIIPGGIISHVATYS	

[illegible][illegible]

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 185596)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beka, F.,
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,
Chapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Deatellano, K., Dewar, K., Domino, M., Doyle, M., Feneator, J.,
Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Landers, T., Lechoczy, J., Levine, R., Lien, C., Liu, G., Locke, K.,
Macdonald, P., Margulis, N., McEwan, P., McGuirk, A., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Olivar, T. M., Peterson, K.,
Pierre, N., Plesani, C., Pollara, V., Raymond, C., Riley, R., Rotman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, W.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 185596)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cooke, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, D. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamaares, R.,
Landers, T., Lechoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnpbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, W.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (03-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 185596)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cooke, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, D. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamaares, R.,
Landers, T., Lechoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnpbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, W.

TITLE
JOURNAL

Submitted (07-MAY-2002) Whitehead Institute/MIT Center for Genome

COMMENT

Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 3, 2002 this sequence version replaced gi:15706183.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: I5662

Center clone name: 709_B_3

FEATURES

source

Location/Qualifiers
1. 185596
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-709B3"
/clone_lib="RP11 Human Male BAC"
1. .381
/rpt_family="L1MA4"
repeat_region
397. .418
/rpt_family="TAAAN"
repeat_region
422. 1327
/rpt_family="MER83B-int"
1328. 1636
/rpt_family="AluSx"
repeat_region
1641. 1953
/rpt_family="AluY"
1954. 1989
/rpt_family="MER83B-int"
1990. 2009
/rpt_family="TAAAN"
repeat_region
2010. 2635
/rpt_family="L1MA4"
repeat_region
2636. 2905
/rpt_family="AluSx"
repeat_region
2906. 3135
/rpt_family="L1MA4"
repeat_region
3136. 3451
/rpt_family="AluY"
repeat_region
3452. 3622
/rpt_family="L1MA4"
repeat_region
3670. 4728
/rpt_family="L1MA4"
repeat_region
4729. 4943
/rpt_family="AluSx"
repeat_region
4945. 5282
/rpt_family="L1MA4"
repeat_region
5283. 5302
/rpt_family="CAA)n"
repeat_region
5303. 5601
/rpt_family="L1MA4"
repeat_region
5602. 5906
/rpt_family="AluSg"
repeat_region
5907. 6220
/rpt_family="L1MA4"
repeat_region
6221. 6526
/rpt_family="AluY"
repeat_region
6527. 7623
/rpt_family="L1MA4"
repeat_region
7624. 7925
/rpt_family="AluSg"
repeat_region
7925. 8104
/rpt_family="L1MA4"
repeat_region
8062. 8068
/note="<30 qual SNGL region"
8196. 8501
/rpt_family="AluSx"
repeat_region

```
repeat_region complement(8556.,.8630)
repeat_region /rpt_family="Tigsters"
repeat_region 8732.,.9247
repeat_region /rptc_family="LM1"
repeat_region 9248.,.9551
repeat_region /rptc_family="AluJo"
repeat_region 9552.,.9716
repeat_region /rptc_family="LM1"
repeat_region 9736.,.9786
repeat_region /rptc_family="LIME"
repeat_region 9787.,.9818
repeat_region /rptc_family="(TG)n"
repeat_region complement(9823.,.10128)
repeat_region /rptc_family="AluSx"
repeat_region 10129.,.10442
repeat_region /rptc_family="LIME"
repeat_region 10469.,.10580
repeat_region /rptc_family="LIME"
repeat_region complement(11726.,.12033)
repeat_region /rptc_family="AluY"
repeat_region 12220.,.12397
repeat_region /rptc_family="MER33"
repeat_region 13586.,.13638
repeat_region /rptc_family="MER33"
repeat_region complement(13639.,.13943)
repeat_region /rptc_family="AluYb"
repeat_region complement(13953.,.14263)
repeat_region /rptc_family="AluSx"
repeat_region 14264.,.14346
repeat_region /rptc_family="MER33"
repeat_region 14443.,.14832
repeat_region /rptc_family="MER45C"
repeat_region 15220.,.15468
repeat_region /rptc_family="MER45B"
repeat_region 15469.,.15608
repeat_region /rptc_family="I2"
repeat_region 15609.,.15984
repeat_region /rptc_family="(TA)n"
repeat_region 16034.,.16224
repeat_region /rptc_family="(TA)n"
repeat_region 16248.,.16441
repeat_region /rptc_family="(TA)n"
```

Query Match	12.6%	Score 103.2	DB 9	Length 185596
Best Local Similarity	54.3%	Pred No. 0.0037		
Matches 253	Conservative	0	Mismatches 208	Indels 5
			Gaps	2
QY	161	ANGTCGAGATATTTTCATCGAAGGCTTAAAGATTAAATGTTATTTGGGATTCTAATTTCT	220	
Db	155139	ATTTTAAATTAATATATATATTTATTTATTAACAATATATTAATATTTATTAATATAA	155198	
QY	221	AGTATTTAATTTGGTGTGGAACTTAAATATATATTAATTCATCCTTGTTTTATTTTTT	280	
Db	155199	TATATTAATATATTTATTTATTAATAATATATATATATTTATTTAATTAATATATATAT	155258	
QY	281	TTTATAACCAATTCACAAATATATTTATTACTTTGGTGTTCATTTATACGGTT	340	
Db	155259	ATATATTTATTTATATATTAATATTAATTAATAATTAATTAATTAATTAATTAAT	155318	
QY	341	AATATTTTACTTAAATATGAGCATTATATATTAATTAATTAATTCATGTGAAAT	400	
Db	155319	AATATTTTAAAAATATATTAATAATTTTAAATTAATAATTTTAAATATATTTAAAT	155378	
QY	401	ATGTAAAAATTAATTTAAAAATATAAATTTATATATATTAATAAAGTCAATCAACAT	460	
Db	155379	ATTATTAATAATTTAAAAATATATATTTTAAAAATATTAATAATTTAAAAATATATAT	155438	
QY	461	AAAAAGATTAATAATCTTAATATATAAATTTTTTAAAAATAGCTTTCAATAATCTGT	520	
Db	155439	TAAATATTTAATAAATATTTAAAAATAT - TATATTTAAAAATTTATATTTAATAATATAT	155499	
QY	521	CAAAACAATAGAAAATATTTTTTGACGGTTCACCAACCGAAAAAGTAATCATTTTC	580	

Db 155498 AAAATATTTTAAATATATATTATATAAATTTTAAAAATATATA----TTTAAAAATATA 155583

QY 581 AGAAAGTAAATCATTTTCAGAAATATATTTTCGGAATTAATTTT 626

Db 155554 TATTTAAATATATATATTTAAAAATTTATATTTAAATTTTATATT 155599

RESULT	12		
AX457067/c			
LOCUS	AX457067	2009 bp	DNA
DEFINITION	Sequence 28 from Patent WO0231186.		linear
ACCESSION	AX457067		
VERSION	AX457067.1	GI:21715849	
KEYWORDS	.		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
	artificial sequences.		

AUTHORS	Beinin, K.
TITLE	Method for the detection of cytosine methylations
JOURNAL	Patent: WO 023186-A 28 18-APR-2002;
FEATURES	Epigenomics AG (DE)
SOURCE	Location/Qualifiers 1. 2009 "archaic construct"

```

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="lower strand of DNA sequence 25, after bisulphite
treatment and P CR"

```

ORIGIN

Query Match	12.6%	Score 102.6	DB 6	Length 2009
Best Local Similarity	49.4%	Pred. No. 0.0013		
Matches 295	Conservative	0	Mismatches 299	Indels 3
				Gaps 1

[illegible]

[illegible]

Query Match	Best Local Similarity	Matches	312; Conservative	0; Mismatches	288; Indels	7; Gaps	3; Indels
139	TGAATAAGTTAAGTGTAGTCTATGCTATGTCGACGAGTATTTTCATGCGAAGGCTTAAGAATTAAT	196					
Db	8654	TAAATATTTATTTATTAATATATTTATTAATTTATTAATATTAATATTTATTTATTTA	8596				
Qy	199	GTTATGGAGTACTAATTTCTAGTATTAATGTGGTTCGACGAGTAAATATAATTAAT	258				
Db	8595	ATTATATATATTAATTAATTAATTAATTTATTAATTTATTAATTTATTAATTAATTAATTAAT	8536				
Qy	259	CAATCCTGTTTATTTATTTTATTTTATTAACAACAATTAACAATAATTTATTAACCTTGG	318				

[illegible]

RESULT 14
 AEO14840/C
 LOCUS AEO14840 249995 bp DNA linear INV 07-OCT-2002
 DEFINITION Plasmodium falciparum 3D7 chromosome 11 section 5 of 8 of the
 complete sequence.
 AEO14840 AEO14186
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 AEO14840 249995 bp DNA linear INV 07-OCT-2002
 Plasmodium falciparum 3D7 chromosome 11 section 5 of 8 of the
 complete sequence.
 AEO14840 AEO14186
 GI:23496193
 Plasmodium falciparum 3D7
 Plasmodium falciparum 3D7
 Eukaryota, Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 1 (bases 1 to 249995)
 Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W.,
 Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T.,
 James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A.,
 Kys, S., Chan, M.-S., Nene, V., Shallow, S.J., Sub, B., Peterson, J.,
 Angiuoli, S., Pertea, M., Allen, J., Selenski, J., Haft, D.,
 Mather, M.W., Vaidya, A.B., Martin, D.M.A., Fairhead, A.H.,
 Frumholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I.,
 Cummings, J.M., Subramanian, G.M., Mungall, C., Venter, J.C., and
 Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and
 Barrell, B.
 Genome sequence of the human malaria parasite Plasmodium falciparum
 Nature 419 (6906), 498-511 (2002)
 12368864
 2 (bases 1 to 249995)
 Gardner, M.J.
 Direct Submission
 Submitted (13-SEP-2002) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 location/Qualifiers
 1..249995
 /organism="Plasmodium falciparum 3D7"
 /mol_type="genomic DNA"
 /isolate="3D7"
 /db_xref="taxon:36329"
 /chromosome="11"

```
repeat_region 108..160
/rpt_type=tandem
repeat_region 418..456
/rpt_type=tandem
repeat_region 540..607
/rpt_type=tandem
repeat_region 996..1020
/rpt_type=tandem
repeat_region 1503..1547
/rpt_type=tandem
repeat_region 1713..1740
/rpt_type=tandem
repeat_region 1946..1984
/rpt_type=tandem
gene
<2062..>2343
/locus_tag="PF11_0269"
<2062..>2343
/locus_tag="PF11_0269"
CDS
2062..2343
/locus_tag="PF11_0269"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAN35853.1"
/db_xref="GI:23496194"
/translation="MDNTEENIFENMTREREKEVLLIENATKREWEISYQWIKRKEPLL
KMLNTHKSHNLQIDVEKCKMGMYNNKYILSCSYNSHVELEMKKYEQS"
repeat_region 2353..2450
/rpt_type=tandem
repeat_region 2353..2447
/rpt_type=tandem
repeat_region 2359..2436
/rpt_type=tandem
repeat_region 2456..2510
/rpt_type=tandem
repeat_region 2619..2658
/rpt_type=tandem
repeat_region 2748..2779
/rpt_type=tandem
repeat_region 2777..2818
/rpt_type=tandem
gene
complement(<2930..>5971)
/locus_tag="PF11_0270"
/locus_tag="PF11_0270"
complement(<2930..>5971)
/locus_tag="PF11_0270"
/codon_start=1
/product="threonine -- tRNA ligase, putative"
/protein_id="AAN35854.1"
/db_xref="GI:23496195"
/translation="MKKILMVIPELLSKICKNNAITLKSLLKKNKIVLANNITTKQEPH
LYHLNNITVNNIYNNNNKKNKNNKNNKIKIKKKENKASFFYINNKKKLFDEQ
GNLILHRRNNWINKRNDIVENYKAVKAKKLLKFLISKNNYIVNTKKEYSACINNNK
TNDVEYMKKLVGEGNEPEIKKLEKKEFKERKQLEENPEELRKRITTEILDGSI
KSGSYVTPDPIALISKRLAEDSIYCKVYVLEKVDVLCDIERGDHNEANDVEE
EKSKEDNNNNNDKNNNDNNNDKNNNDKNNNNNNNNNNNNNNNNNNNNNNNNNNNN
DLAVPLGNCVERFINIQSEBQKIFWSSAHLIGSSLEKLFGEPLTIGPLKGEFY
DIFLNNFSINNEDEKRIEDEPNKLVKNVPEEKYICTREKALELDVNPFLKELIRSK
IPDKKTSVYRCGNFIDLCGPHIKNTGKVFYKVLKSSAYWLGOKENDSLORVYGI
SEOKKSELVEYKLFLEBAKRRDHRVQKLNLFPEEKETSPGSCFWLPHGSKINXLI
EIRKEVIRKREYVLSIPNFSCDLMKTSGHYQKDCMFLFNVENKRWGMPKPNCSG
HCLMFKOLNVSRSIPVRLADPGVILHVEISGSLSGLTRVRFQDDSHIPLCSWEHIK
QEVNLTNLFYVNLFGFKYELFLSTRPKKFIIGQISTWNLAEQHLKDALYAANQWK
INEGDYAGPRIDILVDSQSLNRTHQCGTIQIDFQIPVRFNLQYKNKEVYNNNEK
NEKHNDNDNNNNNNYPINHDDLNHNEKEFKGSONEDKEHNNIHDSLSKNIN
EEGLKKEFERPIIIRHALIGSVRFVAILIEHAGKLPWLSPROAIVLIGPKYND
YANVYETLHNNMPVDIDTSTVNTLAKTRERAGQKQRPYITIVSGEKLITNTVILDR
DDQNNQHYTTIQELINKFNKULDVNSKKNFQIKERNQTTI"
```

```
repeat_region 5733..5794
/rpt_type=tandem
repeat_region 5791..5821
/rpt_type=tandem
repeat_region 5991..6025
/rpt_type=tandem
repeat_region 5993..6123
/rpt_type=tandem
repeat_region 5994..6123
/rpt_type=tandem
repeat_region 6019..6079
/rpt_type=tandem
repeat_region 6078..6125
/rpt_type=tandem
repeat_region 6581..6623
/rpt_type=tandem
repeat_region 6588..6622
/rpt_type=tandem
repeat_region 6924..6961
/rpt_type=tandem
gene
complement(<6962..>11091)
/locus_tag="PF11_0271"
complement(join(<6962..7048,7228..>11091))
/locus_tag="PF11_0271"
complement(join(6962..7048,7228..11091))
/locus_tag="PF11_0271"
/codon_start=1
/product="hypothetical protein, conserved"
/protein_id="AAN35855.1"
/db_xref="GI:23496196"
/translation="MKKKFEENKPSYILKNNNEFKIDISYFTQLHEHKINITYLQ
NYVNLCSSTVYNNKILGFKYKILNRYLIEFAPFLHVATIEINKSPFKYENFDEE
KNNEPNDCTKTIENRNNHINNNDGNKKYOKIYIMNNYNNYGLVILNFTEELF
KNKODHNTVTLPELKYIINNEKNDICOMMLYIHDNIYDPTFEWYKENCILYIEKIN
KYVLSPELDKKYICYSIANPIIKRNNYIKLINSRFRFYIDSKVYINTBRHTI
NIIDIIFYSTKIDTFNNYKQPLNTNIFLLKFPDPIPLHNMDDYDYEYINLKTANI
ECSEDDQSKKEFYQINSFYKLEYLKLDINSQNSHPGNKSFNHNNYNNNSPMHLNY
DMVILFINALSEKEDIKNSKDKILRYIKKDFDLYICFDINYIFNLSWDPFRLAY
CLTLKYLYDFQIDVLAERDLSLRQYVGFESQSEGFIMGYPKVMKSGSINNERY
NDEKKNNDNNYDDKNNNNYDDSHNNNNYDDKNNNNYDDSHNNNNYDDSHNNNNYD
SDLKHIDIMDKNNKSPHYNPINNCLSHQDVSFGSYTMCYNNNSIKDCNDPRDDL
TNEYSHDMPITHEDIHSSQYENNSVNNYTKONRNTKENHNTIHNHLVKTILNS
SLFQVTPDKVHFIYDNGSNYVDINLNKDKDSLNRKQIHILLEKKEGDTGCIINSYK
SPSDEKNDQCIDVASNLGFSINIRKEDNHFTTRVYKQEBMDVLHISEGDENNNNN
ATTNNNNNNIINKYKTFPCDNRKYVDILCGMYVEDKKEKKSIIISILNLPINQDTI
QRISELINIKILKMKILDKPEHILKXILILIGLGLGCMVARNQVWGIOHYTFVD
NSRVSFNSISQYIYTLLEDARKYENIGKYKVAANKNLKICPDLNITAKYMDIPMG
HLNLTNLEDTINELBDLNNHDVVFPLTDSKESRYPCMLTAKQINSLKEIOESV
NNHNNNNNNNNSSSSGSKNFRKGDNVLCSEENRITHEYLENIKCTKIMKSLNNIL
LYEONNNIYKSLNNIHWYDRYOELFYNNILTSVRLCMPLGIVTVALSPDSFVYLRH
SYLFEACAYCNCMHCPSDLSYRTLDEKTVTRCGISNSSISATELIALIQHPL
YEPHPIHDDQYIYNNYDMMQKNSDISNIFSCLAGTPIHNNPNLNFATIKIKIACE
PFECMOCSSEVILIKYQEDKMDPIRNVIRDSIIERITNMQLAVEENDVILIE"
```



```

repeat_region      /rpt_type=tandem
                    11722..11770
repeat_region      /rpt_type=tandem
                    11921..11972
repeat_region      /rpt_type=tandem
                    12249..12278
repeat_region      /rpt_type=tandem
                    12340..12373
gene               /rpt_type=tandem
                    complement(<12386..>13092)
                    /locus_tag="PF11_0272"
mRNA               complement(join(<12386..12652,12889..>13092))
                    /locus_tag="PF11_0272"
CDS                complement(join(12386..12652,12889..13092))
                    /locus_tag="PF11_0272"
                    /codon_start=1
                    /product="ribosomal protein S18, putative"
                    /protein_id="F5AN35856.1"
                    /db_xref="GI:23496197"

```

Query Match 12.5%; Score 102.2; DB 3; Length 24995;
 Best Local Similarity 51.5%; Pred. No. 0.00046;
 Matches 283; Conservative 0; Mismatches 263; Indels 3; Gaps 2;

```

QY 194 TTAATGTTATGGGATTGATTAATTTCTAGATTAATGCTTGAGATTAATATTA 253
    |||
DB 227737 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
    |||
QY 254 TTAATCAATCCTGTTTATTTTATTTTATTAACAATTAATTAATTTTAAAC 313
    |||
DB 227677 ATATATATGTAATTAATTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTT
    |||
QY 314 TTGGTGTGTTCAATTAATGACGTTAATTTTATTTAGTTAATTAATGACATTATTA 373
    |||
DB 227617 TATATTCAGTAATTTGCTGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
    |||
QY 374 TATTAATTAATTAATCATTTGTAATTAATGTAATTAATTAATTAATTAATTTTAA 433
    |||
DB 227557 T-TTATATTAATTAATGCAATGTAATTAATTAATTAATTAATTAATTAATTA
    |||
QY 434 TATATTAATTAATTAATCAATCAATTAATTAATTAATTAATTAATTAATTTTAA 493
    |||
DB 227498 TATATTAATTAATTAATGAGAA--AAAAAATAATTAATTAATTAATTAATTA
    |||
QY 494 TTTAAATAGCTTTCAATTAATCTGTCACAAATTAATTAATTTTTCAGGTTCAATC 553
    |||
DB 227440 ATTTTATCAACAGTTGGAAAAAATAAATAAATAAATAAATAAATAAATAAATA
    |||
QY 554 CAACACCGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 613
    |||
DB 227380 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT
    |||
QY 614 CGAATTAATTTTACTGCAACAAATGAGTCTAGTCTGTTCTGTTTATTTTATTT 673
    |||
DB 227320 TATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT
    |||
QY 674 TTTCTATTTAGAGAACTAGAAATGATTTGCAAAATGCTTTTATCTAGCTTTTAA 733
    |||
DB 227260 ATCCCTATATACATTAATTAATTAATTAATTAATTAATTAATTAATTTTCTT
    |||
QY 734 TTAGTTGAA 742
    |||
DB 227200 TATATTTTAA 227192
    |||

```

RESULT 15
 AC006279/c 245802 bp DNA linear HTG 12-ANG-2000
 LOCUS AC006279
 DEFINITION Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
 PROSCESS ***; 5 unordered pieces.
 AC006279
 VERSION AC006279.8 GI:9797723
 KEYWORDS HTG; HTGS_PHASE1.

SOURCE
 ORGANISM Plasmodium falciparum (malaria parasite P. falciparum)
 Plasmodium falciparum
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE
 1 (bases 1 to 245802)
 AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
 Kurd, O.B., Conway, A.B. and Davis, R.W.
 TITLE Plasmodium falciparum 3D7 chromosome 12
 JOURNAL Unpublished
 REFERENCE
 2 (bases 1 to 245802)
 AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurd, O.B. and Davis, R.W.
 TITLE Direct Submission
 JOURNAL Submitted (05-JAN-1999) Stanford DNA Sequencing and Technology
 Center, Stanford University, 855 California Avenue, Palo Alto, CA
 94304, USA

COMMENT
 On Aug 12, 2000 this sequence version replaced gi:8810451.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 147568: contig of 147568 bp in length
* 147569 147568: gap of unknown length
* 147769 217470: contig of 69702 bp in length
* 217471 217670: gap of unknown length
* 217671 219372: contig of 1702 bp in length
* 219373 219572: gap of unknown length
* 219573 221333: contig of 1761 bp in length
* 221334 221534: gap of unknown length
* 221534 245802: contig of 24269 bp in length.

```

FEATURES
 source
 1..245802
 /organism="Plasmodium falciparum"
 /mol_type="genomic DNA"
 /db_xref="taxon:5833"
 /chromosome="12"
 /clone="PFYA614"
 /clone="3D7"

ORIGIN

Query Match 12.5%; Score 101.8; DB 2; Length 245802;
 Best Local Similarity 50.9%; Pred. No. 0.00051;
 Matches 268; Conservative 0; Mismatches 257; Indels 2; Gaps 1;

```

QY 210 TACTAATTCCTGATTAATTTGCTTGGAAGTTAATTAATTAATTAATTAATTCCTGTT 269
    |||
DB 152433 TATTTTGTGTTAATGTTATGTCGTTAAATTTTATTTATTTATTTATTTATTT
    |||
QY 270 TTTTATTTTCTTTTATTAACAAATTAATTAATTAATTAATTAATTAATTTGCTGTTTCAAT 329
    |||
DB 152373 TTTTCTTTTCTTTTATTAAGAGAGGTAATTAATTAATTAATTAATTAATTAATTA
    |||
QY 330 TTAATGAGGTTAATTTAGTTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 389
    |||
DB 152313 TGTTAATCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
    |||
QY 390 CATTGAAATATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
    |||
DB 152253 CATATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
    |||
QY 450 AATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
    |||
DB 152195 TAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
    |||
QY 510 AGTAATCTGTAACAAATTAATTAATTTTTCAGGTTCAATCAACAAACAGAAAGT 569
    |||
DB 152135 ACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
    |||
QY 570 AATATCTTTCAAGAAAGTAATCAATTTTTCAGAAATTAATTTTTCGAAATTAATTTTACT 629
    |||
DB 152075 TATTTGTAATTAATTAATTTGTTTATTTGTTTGTATTAATTAATTTGATGAT 152016
    |||

```

QY 630 GGCAACAAATGAGCTAGCTGTTCTGTTTATTTTATTTTCTATTTAGAGAA 689
Db 152015 TCAAAATTTATGAGTAAATTTTATTTTATTTTGAATTTTAAATGTC 151956
QY 690 CTAGAAATGATTTGTCAAATGCTTTAATCTAGCTTGTAGATTA 736
Db 151955 ATGTACATTAATCAATATATATTTTATTTTATTTGATATATATA 151909

Search completed: September 20, 2004, 22:34:04
Job time : 3467 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2004, 19:55:10 ; Search time 2676 Seconds

(without alignments)
9105.956 Million cell updates/sec

Title: US-10-089-557-2_COPY_1_816

Sequence: 1 acataggcagcgcgtgc.....tatttcagtgatgtaaaag 816

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estom:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vtl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	194.2	23.8	655	29	CG414267 ZMMBB028
2	178.4	21.9	852	28	BH023559 GH MB000
3	126	15.4	1201	13	BX443774 BX443774
4	120.2	14.7	1201	9	AL536104 AL536104

C	5	119.6	14.7	1056	13	BX415058	BX415058
C	6	118.6	14.5	1101	23	CNS00EVL	AL069706 Drosophila
C	7	118.6	14.5	1200	13	BX436510	BX436510
C	8	118.4	14.5	1200	13	BX415878	BX415878
C	9	117.4	14.4	1201	9	AL536104	AL536104
C	10	116.2	14.2	1101	29	CNS00EVL	AL069706 Drosophila
C	11	116.2	14.2	1201	29	CNS0167M	AL106396 Drosophila
C	12	114.8	14.1	1056	13	BX415058	BX415058
C	13	114.6	14.0	1098	13	BX377526	BX377526
C	14	114.2	14.0	1200	13	BX437758	BX437758
C	15	114.2	14.0	1201	13	BX439779	BX439779
C	16	114	14.0	1200	13	BX415878	BX415878
C	17	112.2	13.8	1101	29	CNS003BD	AL064091 Drosophila
C	18	112	13.7	1201	13	BX430226	BX430226
C	19	111.2	13.6	1076	13	BX38020	BX38020
C	20	110.6	13.5	1201	9	AL565455	AL565455
C	21	109.8	13.5	945	29	CNS04DOK	AL285149 Tetradon
C	22	109.8	13.5	1101	29	CNS00EVL	AL069706 Drosophila
C	23	109.2	13.4	994	29	CNS04DOK	AL285149 Tetradon
C	24	109.2	13.4	1190	29	CNS020N7	AL206908 Tetradon
C	25	108.8	13.3	1163	13	BX462115	BX462115
C	26	108.6	13.3	975	13	BX414927	BX414927
C	27	108.4	13.3	887	13	BX441520	BX441520
C	28	108.4	13.3	1069	29	CX883132	CX883132
C	29	108.4	13.3	1201	13	BX420717	BX420717
C	30	108.4	13.3	1201	13	BX458623	BX458623
C	31	107.6	13.2	1201	13	BX403761	BX403761
C	32	107.6	13.2	1225	29	CNS0161D	AL106171 Drosophila
C	33	107.4	13.2	702	13	BX414899	BX414899
C	34	107.4	13.2	759	29	CNS060XV	AL411257 T7 end of
C	35	107	13.1	1200	13	BX436510	BX436510
C	36	107	13.1	1391	29	CX754863	CX754863
C	37	106.4	13.0	1201	13	BX426629	BX426629
C	38	106.2	13.0	1187	9	AL514085	AL514085
C	39	106	13.0	1200	13	BX437758	BX437758
C	40	105.4	12.9	1200	13	BX456490	BX456490
C	41	105.4	12.9	928	29	CNS00DKY	AL071865 Drosophila
C	42	104.8	12.8	427	13	BX403499	BX403499
C	43	104.8	12.8	1092	29	CNS020K7	AL115696 Tetradon
C	44	104.8	12.8	1165	13	BX38369	BX38369
C	45	104.6	12.8	1101	29	CNS0021J	AL061936 Drosophila

ALIGNMENTS

RESULT 1
CG414267
LOCUS
DEFINITION
ZMMBB0284017.f ZMMBB Zea mays subsp. mays genomic clone
CG414267
ACCESSION
CG414267.1 GI:34504476
VERSION
GSS.
KEYWORDS
Zea mays subsp. mays (maize)
Bukariyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 655)
REFERENCE
Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J. and Wing, R.
Sequencing of the maize genome
Unpublished (2003)
CONTACT: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR PRIMERS

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS00CAP008B02Q01.

FEATURES

Source
1. 1200

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00CAP008B02Q01"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 14.5%; Score 118.4; DB 13; Length 1200;
Best Local Similarity 40.7%; Pred. No. 9.5e-08;
Matches 234; Conservative 75; Mismatches 266; Indels 0; Gaps 0;

```

144 ATGTTAATGAGTCTGATGCTGAGAGTATTTTCATGAGAGGTGTTAAGAAATTAATCTAT 203
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
393 WTTTMMMAADCTTATMTWTHCCATSTTSCCTTTAAMAAATTTBBSSTWATATTTTAT 452
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
204 TGGATTTCTATTTCTGATTAATGCTGAGAGTGAAGTTATATATATATATCAATC 263
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
453 TTTTATTAATTTATTTATTTATTTATTTTAAATTTTAAATTTTAAATTTTAAATTT 512
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
264 CTGTTGTTTATTTTATTTTATTAACAACATTAACAATTAATTTTAACTTGGTGT 323
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
513 WATAWMTTMAATTTTATTTAATTAATTAATTAATTTTAAATTTTAAATTTTAAATTT 572
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
324 TTCAATTTATGAGCGTTAATTTTATTAATTAATTAATTAATTAATTAATTAATTA 383
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
573 AATWMTTWTWTTTWTWMAAAWMTTWTWTTTWTWTTTWTWTTTWTWTTTWTWTTTWT 632
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
384 AATATTCATGTTGAATATATGTAATAATTAATTAATTAATTAATTAATTAATTAAT 443
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
633 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 692
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
444 AAATCAATCAACAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 503
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
693 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 752
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
504 CTTTTCAGTAATCTGTCAACATTAATAATTTTTCAGGTTTCAATCAACACCG 563
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
753 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 812
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
564 AAATGAATCAATTTTCAGAAAGTAATCAATTTTCAGAAATTAATTTTCAGAAATTA 623
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
813 AATTTTAAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 872
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
624 TTTTACGCAACAAATGAGTCTAAGTCTTCTGTTTATTTTATTTTATTTTCAATTA 683
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
873 TAAATTTTAAATTAATTAATTAATTAATTAATTTTAAATTTTAAATTTTAAATTT 932
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
684 GAGAAATCTGAATGATTTGCAATGCTTTAA 718
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
933 TTTWMAATTAATTAATTAATTAATTAATTTTAA 967
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 9
ALS36104 1201 bp mRNA linear EST 31-MAY-2003
LOCUS ALS36104 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS00F022YC18 5-PRIME, mRNA sequence.
ACCESSION ALS36104
VERSION ALS36104.2 GI:31260974
KEYWORDS EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1201)

Li W.B., Gruber C., Jeesee J. and Polayes D.

Full-length cDNA libraries and normalization

Unpublished (2001)

JOURNAL

On Feb 13, 2001 this sequence version replaced gi:1279597.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS00F022B09Q01.

Location/Qualifiers

1. 1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS00F022YC18"

/tissue_type="FETAL BRAIN"

/dev_stage="fetal"

/clone_lib="Homo sapiens FETAL BRAIN"

/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-strand cDNA was digested with Not I and

cloned into the Not I and EcoRV sites of the pCMVSPORT 6

vector. Library was not normalized."

FEATURES

Source

Query Match 14.4%; Score 117.4; DB 9; Length 1201;
Best Local Similarity 37.1%; Pred. No. 1.3e-07;
Matches 194; Conservative 111; Mismatches 217; Indels 1; Gaps 1;

```

193 AATAGTGTATTTGGATTAATTTGATTAATTTGATTAATTTGATTAATTTGATTAAT 252
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
666 AATADTATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 725
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
253 AATATTCATGCTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 311
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
726 AATATTTTWTWTTTWTWTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 785
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
312 AATTGTTGTTTTCATTTATGAGCGTTAATTTTATTTTATTTTATTTTATTTTATTT 371
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
786 AATTTTWTWTTTWTWKAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 845
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
372 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 431
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
846 WATATTTWATTTTWTWMAAAATTTATTTTATTTTATTTTATTTTATTTTATTTATA 905
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
432 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 491
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
906 AATTTTWTWMAATTTTWTWMAATTTTATTTATTTATTTATTTATTTATTTATTTAT 965
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
492 TTTTAAATTAATTTTTCAGTAATCTGTCAACATTAATAATTTTTCAGAGTTCA 551
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
966 AAMWTTTAAWMAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 1025
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
552 TCACAACACGCAAAAGTAATCAATTTTCAGAAAGTAATCAATTTTCAGAAATTAAT 611
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1026 WATWSTATWATTAATAAAWATWATTTTATTAATTAATTAATTAATTAATTAATTA 1085
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
612 TTGCAAAATTAATTTTACGCAACAAATGAGTCTAAGTCTTCTGTTTATTTTATTA 671
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1086 AAATAAAMWATTTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1145
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
672 TTTTCTATTTAGAGAAATCAATTAATTAATTTGCAATGCT 714
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1146 WTTTATADAMDATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1188
    |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```


RESULT 10
CNS00EVL/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION
AL069706
VERSION
AL069706.1 GI:4949849
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuo Ooegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y², cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_1ib="RPCI-98"
/note="end : T7"

ORIGIN
Query Match 14.2%; Score 116.2; DB 29; Length 1101;
Best Local Similarity 37.7%; Pred. No. 2e-07;
Matches 208; Conservative 107; Mismatches 225; Indels 11; Gaps 1;
1;
190 AGAATTAAGTATGGGATTAATTTCTAGTATTAATGTTGGAGTAAAT 249
1006 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 947
250 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 309
946 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 887
310 TAACCTGTTGTTTCAATTAATTAATTAATTAATTAATTAATTAATTA 369
886 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 827
370 TATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 429
826 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 767
430 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 489
766 TATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 707
490 ATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 549
706 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 647

ORIGIN
Query Match 14.2%; Score 116; DB 29; Length 1201;
Best Local Similarity 41.6%; Pred. No. 2e-07;
Matches 226; Conservative 70; Mismatches 260; Indels 1; Gaps 1;
1;
550 CATCCAAACACGAAAGTAAATCAATTTTCAGAAAGTAAATCAATTTTCAGAAATTAAT 609
646 ATT-----AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 598
610 TTTTGAATTAATTTTACTGCGAAACAAATGAGTAAAGTCTGTTCTGTTTATTTT 669
597 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 538
670 TATTTTCAATTAAGAACTAGAAATGATTTGCAATGCTTAATCTACTGTT 729
537 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 478
730 TAGATTAGTTG 740
477 WAAATTTWKKK 467

RESULT 11
CNS0167M
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN15M24 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION
AL106396
VERSION
AL106396.1 GI:5621701
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1201)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
Project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobac11.
FEATURES
source
Location/Qualifiers
1..1201
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN15M24"
/clone_1ib="DrosBAC"
/plasmid="pBelobac11"
/note="end : T7"

ORIGIN
Query Match 14.2%; Score 116; DB 29; Length 1201;
Best Local Similarity 41.6%; Pred. No. 2e-07;
Matches 226; Conservative 70; Mismatches 260; Indels 1; Gaps 1;
1;
127 TTAACCTGTTGGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 186
632 TTTAGCTTAAKKTGKTTTBTBTTTGTGKKKKTGKKTGTTTTGTTTTGTTTTGTTTTGTTTT 691
187 TTAAGATTAATGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 246
692 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 751
247 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 305
752 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 811

QY 306 TATTACCTTGGTGGTTTCATTTAAGACGTTAATATTTAGTTAAATTTGACGA 365
 DB 812 TTTTAAAAAAATTTATATATATATATATATATATATATATATATATATAT 871
 QY 366 TTTAT 425
 DB 872 TAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 931
 QY 426 TTTTAT 485
 DB 932 AATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 991
 QY 486 AATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 545
 DB 992 TTTATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1051
 QY 546 GGTTCATCCAAACACGAGAAAGTAAATCATTTTCAGAAAAGTAAATCATTTTCA 605
 DB 1052 AAAAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1111
 QY 606 TTTATTTTGGAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 665
 DB 1112 TTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1171
 QY 666 TTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 692
 DB 1172 TTTATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1198

RESULT 12
 LOCUS BX415058 1056 bp. mRNA linear EST 15-MAY-2003
 DEFINITION BX415058 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP004YG19
 3-PRIME, mRNA sequence.
 ACCESSION BX415058
 VERSION BX415058.1 GI:30767520
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1056)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0CAP004AD10NP1.

FEATURES
 source
 1..1056
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0CAP004YG19"
 /tissue_type="THYMUS"
 /clone_lib="Homo sapiens THYMUS"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with NotI and cloned into
 the NotI and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN
 Query Match 14.1%; Score 114.8; DB 13; Length 1056;
 Best Local Similarity 38.1%; Pred. No. 3.3e-07;
 Matches 225; Conservative 100; Mismatches 265; Indels 1; Gaps 1;
 184 GTGTTAAGATTATGATTGGAATTACTAATTTCTAGTATTAATTTGGTTGGAGT 243

DB 333 GTTATTAACCTGGTGGTCTGCAATATTTATATATATATATATATATATATAT 392
 QY 244 TAAAT 303
 DB 393 AAAAAAAT 452
 QY 304 TTTATTTACCTGGTGGTGGTCTGCAATATTTATATATATATATATATATATAT 363
 DB 453 TATATTTAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 512
 QY 364 CAT 422
 DB 513 WDKKKKKTAT 572
 QY 423 AATTTAT 482
 DB 573 AAAAAAAT 632
 QY 483 TATATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 542
 DB 633 WATATTTTAT 692
 QY 543 GCAGGTTCATCCAAACACGAGAAAGTAAATCATTTTCAGAAAAGTAAATCATTTTCA 602
 DB 693 KATKAKTKTAT 752
 QY 603 AATTTATTTTGGAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 662
 DB 753 AAAAAAAT 812
 QY 663 TATTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 722
 DB 813 TTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 872
 QY 723 GCTGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 773
 DB 873 TAT 923

RESULT 13
 LOCUS BX377526/c 1098 bp. mRNA linear EST 08-MAY-2003
 DEFINITION BX377526 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1007YB08 5-PRIME, mRNA sequence.
 ACCESSION BX377526
 VERSION BX377526.1 GI:30460084
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1098)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0D1007DA04QPL.
 FEATURES
 source
 1..1098
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1007YB08"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)

ORIGIN

primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match

14.0%; Score 114.6; DB 13; Length 1098;

Best Local Similarity 41.4%; Pred. No. 3.4e-07; Matches 272; Conservative 78; Mismatches 305; Indels 2; Gaps 2;

```
QY 161 ATGCGAGATATTTTCAGAGCTGTTAAGATTAATGTTATGGATTAATTTCT 220
DB 1047 ATGAGAGATATTTTCAGAGCTGTTAAGATTAATGTTATGGATTAATTTCT 220
QY 221 AGTATTAATGTTGTTGAGGTTAATATATATATATATATATATATATATAT 280
DB 987 TTTATTAATGTTGTTGAGGTTAATATATATATATATATATATATATATAT 280
QY 281 TTTATTAATGTTGTTGAGGTTAATATATATATATATATATATATATATAT 339
DB 927 TATATTAATGTTGTTGAGGTTAATATATATATATATATATATATATATAT 339
QY 340 TATATTAATGTTGTTGAGGTTAATATATATATATATATATATATATATAT 399
DB 867 TATATTAATGTTGTTGAGGTTAATATATATATATATATATATATATATAT 399
QY 400 TATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 459
DB 807 TATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 459
QY 460 TATAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 519
DB 747 TATAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 519
QY 520 TCAAGCA-ATGAGAAATATTTTTCAGGTTCCACACACAGAAAGTAACTTT 578
DB 687 TCAAGCA-ATGAGAAATATTTTTCAGGTTCCACACACAGAAAGTAACTTT 578
QY 579 TCAGAAAGTAAATTCATTTTTCAGAAATTAATTTTTCAGAAATTAATTTTTC 638
DB 627 TCAGAAAGTAAATTCATTTTTCAGAAATTAATTTTTCAGAAATTAATTTTTC 638
QY 639 ATGAGCTTAAGTCTTCTGTTTATTTTATTTTATTTTATTTTATTTTATTT 698
DB 567 ATGAGCTTAAGTCTTCTGTTTATTTTATTTTATTTTATTTTATTTTATTT 698
QY 699 GATTTGCTCAATCTCTTAATCTAGCTGTTAGATTAAGTGAAGGACAGAA 758
DB 507 GATTTGCTCAATCTCTTAATCTAGCTGTTAGATTAAGTGAAGGACAGAA 758
QY 759 TTGTCAGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 815
DB 447 TTGTCAGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 815
```

RESULT 14
BX437758/c
LOCUS BX437758 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YB01
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX437758
VERSION BX437758.1 GI:30773605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France

FEATURES
source

Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0CAP008CA01QPL
Location/Qualifiers
1.1200
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP008YB01"
/issue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and EcoRV into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match

14.0%; Score 114.2; DB 13; Length 1200;

Best Local Similarity 34.5%; Pred. No. 3.6e-07; Matches 201; Conservative 120; Mismatches 262; Indels 0; Gaps 0;

```
QY 127 TTAAGCTTTGCGAAATGTTAATGCTAGCTAGCTAGCTAGCTAGCTAGCT 186
DB 1029 TTAAGCTTTGCGAAATGTTAATGCTAGCTAGCTAGCTAGCTAGCTAGCT 186
QY 187 TTAAGCTTTGCGAAATGTTAATGCTAGCTAGCTAGCTAGCTAGCTAGCT 246
DB 969 TTAAGCTTTGCGAAATGTTAATGCTAGCTAGCTAGCTAGCTAGCTAGCT 246
QY 247 TTAAGCTTTGCGAAATGTTAATGCTAGCTAGCTAGCTAGCTAGCTAGCT 306
DB 909 TTAAGCTTTGCGAAATGTTAATGCTAGCTAGCTAGCTAGCTAGCTAGCT 306
QY 307 ATTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
DB 849 ATTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
QY 367 ATTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
DB 789 ATTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
QY 427 ATTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
DB 729 ATTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
QY 487 TTAATTTTAAATATGCTTTTCAGTAAATCTGCAACATATGAAATATTTT 546
DB 669 TTAATTTTAAATATGCTTTTCAGTAAATCTGCAACATATGAAATATTTT 546
QY 547 GTTCAATCAACACAGAAAGTAAATCTGTTTCAGAAAGTAAATCTGTTTC 606
DB 609 GTTCAATCAACACAGAAAGTAAATCTGTTTCAGAAAGTAAATCTGTTTC 606
QY 607 TATTTTGGAAATATTTTTCAGTAAATCTGTTTCAGAAAGTAAATCTGTT 666
DB 549 TATTTTGGAAATATTTTTCAGTAAATCTGTTTCAGAAAGTAAATCTGTT 666
QY 667 TTTTATTTTTCATTTTGAAGAACTGAAATGATTTGCTGCTGCTGCT 709
DB 489 TTTTATTTTTCATTTTGAAGAACTGAAATGATTTGCTGCTGCTGCT 709
```

RESULT 15
BX439779/c
LOCUS BX439779 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE014YF05
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION BX439779
VERSION BX439779.1 GI:30771778
KEYWORDS EST.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2004, 20:51:30 ; Search time 472 Seconds
(without alignments)
8726.094 Million cell updates/sec

Title: US-10-089-557-2_COPY_1_816

Sequence: 1 actataggacgacgctgctc.....tatttcagctatgtaaaag 816

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124.2	15.2	606	US-10-021-323-7696	Sequence 7696, App
2	92.6	11.3	520	US-10-021-323-7699	Sequence 7699, App
3	92.4	11.3	3673778	US-10-312-841-1	Sequence 1, Appl
4	90.6	11.1	469	US-10-021-323-16830	Sequence 16830, A
5	90.2	11.1	7351	US-10-311-455-2	Sequence 2, Appl
6	90.2	11.1	8305	US-10-311-455-1542	Sequence 1542, App
7	89.8	11.0	18357	US-10-433-793-114	Sequence 114, App
8	89.4	11.0	6106	US-10-221-714A-151	Sequence 151, App
9	89.4	11.0	6106	US-10-311-455-1445	Sequence 1445, App
10	89.4	11.0	6106	US-10-257-166-113	Sequence 113, App
11	89.4	11.0	158001	US-10-211-179-11	GENERAL INFORMATI
12	89.2	10.9	565	US-10-021-323-11125	Sequence 11125, A
13	88.8	10.9	520	US-10-021-323-7699	Sequence 7699, App
14	88.2	10.8	560	US-10-021-323-2253	Sequence 2253, App

15	88.2	10.8	16602	US-10-311-455-700	Sequence 700, App
16	87.8	10.8	419	US-09-960-352-11234	Sequence 11234, A
17	87.4	10.7	17848	US-10-239-676-28	Sequence 28, Appl
18	87.4	10.7	17848	US-10-240-453-38	Sequence 38, Appl
19	87.4	10.7	17848	US-10-257-166-58	Sequence 58, Appl
20	87.2	10.7	15548	US-10-311-455-2128	Sequence 2128, App
21	86.6	10.6	158001	US-10-211-179-11	GENERAL INFORMATI
22	86	10.5	38342	US-10-221-714A-472	Sequence 472, App
23	85.8	10.5	469	US-10-021-323-16830	Sequence 16830, A
24	85.8	10.5	5487	US-10-311-455-1571	Sequence 1571, App
25	85.2	10.4	543	US-10-021-323-12208	Sequence 12208, A
26	85.2	10.4	6465	US-10-311-455-958	Sequence 958, App
27	85	10.4	3673778	US-10-312-841-2	Sequence 2, Appl
28	84.8	10.4	5845	US-10-311-455-1636	Sequence 1636, App
29	84.6	10.4	27890	US-10-741-601-5686	Sequence 5686, App
30	84.6	10.4	126872	US-10-741-601-5738	Sequence 5738, App
31	84.4	10.3	6063	US-10-221-714A-59	Sequence 59, Appl
32	84.4	10.3	6216	US-10-257-166-13	Sequence 13, Appl
33	84.4	10.3	13511	US-10-311-455-254	Sequence 254, App
34	84.2	10.3	6478	US-10-239-676-144	Sequence 124, App
35	84.2	10.3	6478	US-10-240-453-144	Sequence 144, App
36	84.2	10.3	6626	US-10-221-714A-536	Sequence 536, App
37	84	10.3	960	US-10-198-846-6381	Sequence 6381, App
38	84	10.3	6881	US-10-311-455-1353	Sequence 1353, App
39	84	10.3	6881	US-10-240-453-123	Sequence 123, App
40	83.8	10.3	560	US-10-021-323-2253	Sequence 2253, App
41	83.6	10.2	7008	US-10-221-714A-254	Sequence 254, App
42	83.4	10.2	56153	US-10-221-714A-520	Sequence 520, App
43	83.4	10.2	3673778	US-10-312-841-2	Sequence 2, Appl
44	83.2	10.2	6106	US-10-221-714A-152	Sequence 152, App
45	83.2	10.2	6106	US-10-311-455-1446	Sequence 1446, App

ALIGNMENTS

RESULT 1
US-10-021-323-7696/c
Sequence 7696, Application US/10021323
GENERAL INFORMATION:
Application No. US20040123340A1
APPLICANT: Deikman, Jill
APPLICANT: Feng, Paul C.C.
APPLICANT: Fincher, Karen L.
APPLICANT: Ziegler, Todd E.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21 (52274) B
CURRENT APPLICATION NUMBER: US/10/021,323
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255, 619
PRIOR FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 17880
SEQ ID NO 7696
LENGTH: 606
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
NAME/KEY: unsure
LOCATION: (1).. (606)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: LIB5828-019-Q1-N6-D2
US-10-021-323-7696

Query Match 15.2%; Score 124.2; DB 17; Length 606;

Best Local Similarity 82.9%; Pred. No. 7.6e-08;

Matches 155; Conservative 0; Mismatches 28; Indels 4; Gaps 1;

QY 186 GTTAAAGATTATGATTTGCGATTAATTTCTAGATTAATTTGGTTGGAAGTTA 245
|||
Db 183 GTCAACATTTCCAGCGGTTTATTATTTATTTCTAGATTAATTTGGTTGGAAGTTA 124
|||
QY 246 ATATATATATTTCAATCCCTGTTTATTTTATTTTATTTTATTAACAACATTAATTAAT 305

Db 123 ATTATATATATATCAGTCC---TTTTTTTATTTTATTAACCAATTTAAATATTT 68
QY 306 TATTAACTTTGGTGTTCATTTATGACGGTAAATTTTATTTAAATATGAGCA 365
Db 67 TATTGGCTTTGGTGTTCATTTATGACGGTAAATTTTATTTATTTATTTAGTA 8
QY 366 TTATTTAT 372
Db 7 TTACTAT 1

RESULT 2

US-10-021-323-7699
; Sequence 7699, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 7699
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LHB828-019-Q1-N6-D6
US-10-021-323-7699

Query Match 11.3%; Score 92.6; DB 17; Length 520;
Best Local Similarity 53.1%; Pred. No. 0.0013;
Matches 197; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 209 TTACAAATTTTCAGATTAATTTGCTTGGAACTTAATATATTTATTCATCTCTGT 268
Db 13 TTTTAAATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 72
QY 269 TTTTAAATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 328
Db 73 TTTTAAATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 132
QY 329 TTTATGACGGTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 388
Db 133 TTTTAAATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 192
QY 389 TCATTTGTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 448
Db 193 AAAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 252
QY 449 CAATCAAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 508
Db 253 AAAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 312
QY 509 CAGTAATCTGTCAAAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTT 568
Db 313 AAAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 372
QY 569 TAAATCATTTT 579
Db 373 AAAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT

RESULT 3
US-10-312-841-1
; Sequence 1, Application US/10312841

; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenetics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 11.3%; Score 92.4; DB 15; Length 3673778;
Best Local Similarity 48.9%; Pred. No. 0.016;
Matches 275; Conservative 0; Mismatches 286; Indels 1; Gaps 1;

QY 145 TGTATATGTAGTGTGTATGAGATTTTTCATGAGAGGTATTAAGATTTAAATTTT 204
Db 205750 TTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 205809
QY 205 GGAATTAATTTCTAGATTAATTTGCTTGGAACTTAATTAATTTATTTATTTATTT 264
Db 205810 TTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 205869
QY 265 TTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 324
Db 205870 TTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 205929
QY 325 TCAATTTATGACGGTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTT 384
Db 205930 TTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 205989
QY 385 TAAATCATTTGTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 443
Db 205990 TTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 206049
QY 444 AAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 503
Db 206050 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 206109
QY 504 CTTTCAATTAATTTGCTTCAATTAATTTGCTTCAATTAATTTGCTTCAATTAATTT 563
Db 206110 TTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 206169
QY 564 AAATCAATCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTT 623
Db 206170 TTTTAAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 206229
QY 624 TTTTCAATTAATTTGCTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCA 683
Db 206230 GTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 206289
QY 684 GAGAAATCAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 705
Db 206290 GCGGCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 206311

RESULT 4
US-10-021-323-16830/C
; Sequence 16830, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.

```

RESULT 6
US-10-311-455-1542
; Sequence 1542, Application US/10311455
; Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detecting Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1542
; LENGTH: 8305
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1542

```

Query Match 11.1%; Score 90.2; DB 15; Length 8305;
Best Local Similarity 50.1%; Pred. No. 0.0058;
Matches 284; Conservative 0; Mismatches 273; Indels 10; Gaps 2;

QY 173 TTTCTGAGAGGTGTTAAGATTTATGTTATGGATCTACTAATTTCTAGATTTAATGTT 232
DB 1710 TTTTATGTTGGTAAATTAAGATTTGTTTATTTTATATATATATATATATATAGCT 1769
QY 233 GGTGGAAGTAAAT 292
DB 1770 ATATAGAT 1829
QY 293 ATTACAAATATTTATTTTACTGTTGTTTCAATTTATGACGCTTAATATTTTACTT 352
DB 1830 AT 1889
QY 353 TATATATGAGCATTTAT 412
DB 1890 AT 1949
QY 413 TTTAAAT 472
DB 1950 AT 2003
QY 473 AATCTTAAAT 532
DB 2004 AT 2063
QY 533 AATATTTTGGAGGTGATCCAAACACGAAAGTAATGCTTTGAGAAATTAAT 592
DB 2064 TAT 2119
QY 593 CATTTTGGAAATATTTTTCGAAATATTTTACGCAACAAATGAGCTTAAGT 652
DB 2120 TAT 2179
QY 653 TTTCTGTTTATTTTATTTTCTATTTAGAGAACTGAGAAATGATTTGCTCAATGT 712
DB 2180 TAT 2239
QY 713 CTTTATCTAGCTGTTTATAGTTAGTT 739
DB 2240 ATATTTATTTAGAGCTTTGAGATTTAGTT 2266

RESULT 7
US-10-433-793-114
; Sequence 114, Application US/10433793
; Publication No. US20040142334A1
; GENERAL INFORMATION:
; APPLICANT: Edigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433,793
; PRIORITY FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 114
; LENGTH: 18357
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-114

Query Match 11.0%; Score 89.8; DB 17; Length 18357;
Best Local Similarity 48.9%; Pred. No. 0.0081;
Matches 302; Conservative 0; Mismatches 307; Indels 8; Gaps 2;

QY 124 GCATTACTGTTGGTGAATATGTTAATGCTAGTCTATGCGAAGTATTTTCATGGAAG 183
DB 15286 GATTTAATTTATTTATAGTGTGTTAGTAGTGTGCGGAGTGTGTTTATAGAGATGCT 15345
QY 184 GTGTTAAGATTTAATGTTATTTGGGATTTACTAATTTCTAGATTTATTTGCTTGGAGCT 243

DB 15346 TTTAT 15405
QY 244 TAT 303
DB 15406 TTAATGCGAAGTATTTTATTTGTTTATTTGTTTATTTGTTTATTTGTTTATTTGTT 15465
QY 304 TTTATTTTACTGTTGTTGTTTCAATTTATGACGCTTAATATTTTATGTTTATATATGAG 363
DB 15466 TTTTATTTATTTGTTTAT 15525
QY 364 CATTTAT 423
DB 15526 GTGTTTATGAAATTTTAAAGAGATTTGCTAAATATATATATATATATATATATAT 15585
QY 424 AATTTAT 483
DB 15586 GATTTATTTTAAATGTTTATTTTATTTTATATATATATATATATATATATATATAT 15645
QY 484 ATATATATATTTTAT 543
DB 15646 AATATATGTTTAT 15698
QY 544 CAGGTTATCCAAACACGAAAGTAATCATTTTTCGAAAGTAATCATTTTTCGAA 603
DB 15699 TATGTTATTTTAT 15758
QY 604 AATTTTTCGAAATATTTTATTTTACGCAACAAATGAGCTTAAGTGTGCTTTT 662
DB 15759 TTTTATGTTTATTTTAT 15818
QY 663 TATTTTATTTTCTATTTTATGAGAACTGAGAAATGATTTGCTCAATGCTTTAATCTA 722
DB 15819 TATTTTGGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15878
QY 723 GCTTGTATAGTTAGTT 739
DB 15879 TTTTATTTAGTTTATTTT 15895

RESULT 8
US-10-221-714A-151
; Sequence 151, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; PRIORITY FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 151
; LENGTH: 6106
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-151

Query Match 11.0%; Score 89.4; DB 13; Length 6106;
Best Local Similarity 51.7%; Pred. No. 0.0068;
Matches 278; Conservative 0; Mismatches 253; Indels 7; Gaps 3;

QY 198 TGTATTTGGATTACTAATTTCTAGTATTAATGCGTTTGAAGTTAATATATATAT 257
DB 2612 TTTTATTTTAATGTTTTGTTGTTTAAAGTTAAAGAAATATATTAATAGTTA 2671

QY 258 TCAATCCTGTTTTTATTTTTTTTATACACAAATTAACAAATTTAATTAACCTTG 317
DB 2672 TTAATGTTTAAATTTAGTTTTTAAATATATATATATATATATATATATATAGTTA 2731

QY 318 GTTGTTCATTTATGACGGTTAATTTTAGTTAATATGAGCATTTATATAT 377
DB 2732 TAT 2791

QY 378 AAT 437
DB 2792 AAT 2850

QY 438 TAT 492
DB 2851 TAT 2910

QY 493 TTTTAAATAGCTTTTCAGTAAATCTGCAACATAGAAATATTTTTCAGAGTTTCA 552
DB 2911 TTTTAAATAGCTTTTCAGTAAATCTGCAACATAGAAATATTTTTCAGAGTTTCA 2970

QY 553 CCAACACGAGAAAGTAAATCTGTTTCAGAAAGTAAATCTGTTTCAGAAATATTTT 612
DB 2971 -TAT 3029

QY 613 TCGAATATATTTTACTGCGAAACAAAGAGCTAGTGTCTGTTTTTATTTTAT 672
DB 3030 ATATTAAT 3089

QY 673 TTTTCATTTAGAGAAATAGAAATGATTTGCAATGCTTAAATCTGCTGTTT 730
DB 3090 TATAGATTTTAT 3147

RESULT 9
US-10-311-455-1445
; Sequence 1445, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation of Cytosine
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1445
; LENGTH: 6106
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1445

Query Match 11.0%; Score 89.4; DB 15; Length 6106;
Best Local Similarity 51.7%; Pred. No. 0.0068;
Matches 278; Conservative 0; Mismatches 253; Indels 7; Gaps 3;

QY 198 TGTATTTGGATTACTAATTTCTAGTATTAATGCGTTTGAAGTTAATATATATAT 257
DB 2612 TTTTATTTTAATGTTTTGTTGTTTAAAGTTAAAGAAATATATTAATAGTTA 2671

QY 258 TCAATCCTGTTTTTATTTTTTTTATACACAAATTAACAAATTTAATTAACCTTG 317
DB 2672 TTAATGTTTAAATTTAGTTTTTAAATATATATATATATATATATATATATAGTTA 2731

QY 318 GTTGTTCATTTATGACGGTTAATTTTAGTTAATATGAGCATTTATATAT 377
DB 2732 TAT 2791

QY 378 AAT 437
DB 2792 AAT 2850

QY 438 TAT 492
DB 2851 TAT 2910

QY 493 TTTTAAATAGCTTTTCAGTAAATCTGCAACATAGAAATATTTTTCAGAGTTTCA 552
DB 2911 TTTTAAATAGCTTTTCAGTAAATCTGCAACATAGAAATATTTTTCAGAGTTTCA 2970

QY 553 CCAACACGAGAAAGTAAATCTGTTTCAGAAAGTAAATCTGTTTCAGAAATATTTT 612
DB 2971 -TAT 3029

QY 613 TCGAATATATTTTACTGCGAAACAAAGAGCTAGTGTCTGTTTTTATTTTAT 672
DB 3030 ATATTAAT 3089

QY 673 TTTTCATTTAGAGAAATAGAAATGATTTGCAATGCTTAAATCTGCTGTTT 730
DB 3090 TATAGATTTTAT 3147

RESULT 10
US-10-257-166-113
; Sequence 113, Application US/10257166
; Publication No. US20040023230A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of Cytosine
; FILE REFERENCE: 5013.1011
; CURRENT APPLICATION NUMBER: US/10/257,166
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: PCT/EP01/07470
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 113
; LENGTH: 6106
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-113

Query Match 11.0%; Score 89.4; DB 17; Length 6106;
Best Local Similarity 51.7%; Pred. No. 0.0068;
Matches 278; Conservative 0; Mismatches 253; Indels 7; Gaps 3;

QY 198 TGTATTTGGATTACTAATTTCTAGTATTAATGCGTTTGAAGTTAATATATATAT 257
DB 2612 TTTTATTTTAATGTTTTGTTGTTTAAAGTTAAAGAAATATATTAATAGTTA 2671


```
RESULT 13
US-10-021-323-7699/C
; Sequence 7699, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Ziegler, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 7699
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3828-019-Q1-N6-D6
US-10-021-323-7699

Query Match      10.9%; Score 88.8; DB 17; Length 520;
Best Local Similarity 50.4%; Pred. No. 0.0042;
Matches 243; Conservative 0; Mismatches 237; Indels 2; Gaps 1;

QY 127 TTAACGTTGGGAAAGTATGATGCTATGCTATGCAAGATTTTCATGAGAGG 186
DB 485 TTTAAATTTTATTTTATTTTATTTTATTTTAAATAATTAATTTTATTTT 426
QY 187 TTAAGATTATGTTATGAGATTACTAATTTCTAGTATTAATGTTGGAGTTAA 246
DB 425 TTCAATCATCTTTTATTTTATTTTATTTTCAAAAAAATATATTTTATTTT 366
QY 247 TATATATTTATTCATCTGTTTATTTTATTTTATTTTATACACAATTCAGATTA 306
DB 365 TTTTATTTTAAATAATTTTATTTTATTTTATTTTAAATAATTTTAAAAAT 306
QY 307 ATTACCTTGGTGGTTTCATTTATGACGTTAATTTTATGTTAATTAATGACAT 366
DB 305 TTTTATTTTATCTCTTTTATTTTATTTTAAATTAATCTTTTATTTTATTT 246
QY 367 TATTATATATTAATAATAATCATGTAATATATATGTAATAATTAATTAATTAAT 426
DB 245 TTTTCTTTTAA- AATTTACCTTTTATTTTATTTTAAATTTTATTTTATTTT 188
QY 427 TTTATATATATATTAATACTCAATCAACAATAAAGATTAATTTCTTAATTA 486
DB 187 TTTATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 128
QY 487 TAAATTTTAAATAGCTTTTCAGTAATACTGCAACAATGAGAAATATTTTGGAG 546
DB 127 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 68
QY 547 GTTCATCCAAACACAGAAAGTAATCAATTTTCAGAAAAAGTAATCAATTTTCAGAAAT 606
DB 67 AAAAAAAAAAAAAAAAAAAAAAAAAAAGAAACAAAGAAATAAATAAATAAATAAATAAAT 8
QY 607 TA 608
DB 7 AA 6

RESULT 14
US-10-021-323-2253
; Sequence 2253, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
```

```
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 2253
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3825-011-Q1-N6-C9
US-10-021-323-2253

Query Match      10.8%; Score 88.2; DB 17; Length 560;
Best Local Similarity 51.1%; Pred. No. 0.0051;
Matches 207; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 209 TTAATAATTTCTAGATTAATGTTGGTTGGAAGTTAATATATTAATTAATCAATCCCTGT 268
DB 6 TTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 65
QY 269 TTTTATTTTTTTTTTTTAAACACAATTCAGAAATTAATTAATTAATCTGTTGTTTCAA 328
DB 66 TTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 125
QY 329 TTTATACGCTTAATTTTATGTTAATATGAGATTAATTAATTAATAATAATAA 368
DB 126 TTTTATTTTTTTTTTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTAAAAA 185
QY 389 TCATGTAATATATGTAATAATTAATTAATAATTAATTAATTAATTAATTAATTAAT 448
DB 186 TAAATTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 245
QY 449 CAATCAACATATAAAGATAATTAATCTTAATTAATTAATTAATTAATTAATTAATTAAT 508
DB 246 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 305
QY 509 CAGTAATCTGCAACAATAGAAATATTTTGGAGGTTCTCAACACACAGAAAG 568
DB 306 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGTTAAAAA 365
QY 569 TAAATCATTTTCAGAAAGTAATCAATTTTTCAGAAATTAATTTT 613
DB 366 AAAAAAAAAATTTTAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 410

RESULT 15
US-10-311-455-700
; Sequence 700, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 700
; LENGTH: 16602
```

